THE PROCEEDINGS OF
THE PHYSICAL SOCIETY

Section B

VOL. 62, PART 3 — 1 March 1949 — No. 351 B

A Photoelectric Type of Acoustic Spectrograph using Sound Film

BY D. BROWN, C. F. COLEMAN AND J. W. LYTTLETON
Auckland University College, New Zealand

Communicated by P. W. Burbidge; MS. received 30th March 1948

ABSTRACT. An account is given of an apparatus which analyses non-recurrent waveforms such as occur in speech, so that frequency-distribution is represented against time in an FT-diagram. From the theory governing such diagrams, no unique representation is possible, the result depending on the analysing convention adopted, but all conventions are governed by the uncertainty relation $\Delta F \Delta T \geqslant 1$. The method employed is an optical one which analyses a continuously moving sound film by means of a frequency-scanning disc, so producing an oscillograph pattern which is photographed on to another moving film. Typical analyses are reproduced and discussed, mainly "speech patterns" which show the composition very clearly. Advantages include the ability to record the amplitude and phase of each component, also the ability to change the analysing convention in any desired manner, for example to give the error-function convention which leads to minimum uncertainty. The paper concludes with considerations of resolving power, spectral line profiles, and amplifier band width requirements.

§1. INTRODUCTION

IN this paper, of which a brief summary has already been published (Brown and Lyttleton 1947), an account is given of a practical method for obtaining a continuous registration of the spectra of disturbances such as speech sounds. It has hitherto been difficult to determine the composition of aperiodic disturbances, particularly those which vary rapidly with time, and it is only recently that instrumental methods have been evolved for performing the analysis rapidly. The equivalent mathematical representation of such disturbances in terms of both frequency f and time t, intermediate between the conventional representations in terms of f or t alone, has recently been worked out by Gabor (1946). This representation can be depicted by an "information diagram" in which f, t appear as Cartesian coordinates, and amplitude is represented by, say, photographic density. With Gabor's definitions of Δf and Δt, there is an uncertainty product governed by the relation $\Delta f \Delta t \geqslant \frac{1}{2}$ and represented in the diagram by the weighted area occupied by the signal.

Instruments designed to produce analyses of the above type may for convenience be called FT-analysers, to distinguish them from the extreme cases represented by T-analysers (oscillographs) and F-analysers (spectrographs). In a previous paper (Brown 1939) it was shown that FT-analyses could be obtained by an optical diffraction method from a sound film record of the disturbance. While this method is simple and theoretically satisfying, and has led to some

practical applications (Schouten 1939), it is not very satisfactory when used with ordinary commercial film, which is liable to contain small optical irregularities to which the diffraction process is unfortunately rather sensitive.

The method described in the present paper again employs sound film recording, and is the outcome of a search for a method which would be unaffected by the small film irregularities referred to above. The method, which was devised in the early war years, is a photoelectric one and produces the desired spectra by means of a special frequency-scanning disc. The principle of this disc was also evolved independently by Fürth and Pringle (1946) for use in a Fourier transformer for aperiodic functions, these being represented by appropriate masks or silhouettes; on the above classification their instrument is thus an F-analyser. Our own account of the method has been delayed by preoccupation with the larger issues involved in using the disc-principle to produce FT-representations, and the necessity for holding over practical development till after the war.

The earliest and best known approach to the problem of FT-analysis is the one using banks of band-pass filters. During the present investigations, a series of articles has appeared describing recent developments in the Bell Telephone Laboratories (Potter *et al.* 1946), in which this method is used in a refined form to give what are called "visible speech translators". Another article in the series describes a different type of FT-analyser using magnetic recording in the initial stage, the subsequent analysis being made by a heterodyne circuit with a fixed band-pass filter. It is with this last method, for which a great deal of development work is stated to have been necessary, that the greatest number of acoustic spectra have been investigated to date.*

The apparatus described in the present paper must be regarded as in a less advanced stage of development, circumstances having stood in the way of intensive development work. However, the various alternative methods with their different individual features may ultimately be found to supplement one another. The present method offers the following advantages which should make its further development desirable: (i) the T-resolution can be easily and smoothly varied to suit the particular type of disturbance it is wished to analyse; (ii) the disturbances are analysed in their "natural" sequence, i.e. the FT-diagram is built up by a unidirectional motion along the T-axis associated with a "repeated-F scan"; thus there is no limit to the length of film which can be analysed in a single continuous operation; (iii) the adoption of the natural sequence makes it a simple matter to stop the recorder at any particular stage of the disturbance and to obtain a quantitative record of the amplitudes of the F-components present in the region concerned; (iv) it offers a practical means of investigating and recording the phases of the components; (v) the shape of the T-response curve can be readily varied, so as to give for example a probability curve, which Gabor (1946) has shown to be the condition for minimum uncertainty.

§ 2. EXPERIMENTAL ARRANGEMENTS

(i) *Recording*

While the present method can be used with slight modification to analyse either variable-density or variable-area sound film, the former type is easier

* An account of the Bell Telephone Laboratories' investigations is now available in book form: *Visible Speech*, by R. K. Potter, G. A. Kopp and H. C. Green (publ. D. Van Nostrand Co.)

to make and to analyse and has so far been used exclusively. The process of making such films is a fairly standard one and needs little description. The apparatus used in this case was built round a Cossor recording camera, suitably modified, a gas discharge recording tube being used which produced a sound-track 5 mm. wide along the centre of the 35 mm. film. The microphone response was sensibly flat throughout the audio-frequency range, but the amplifier was arranged to give a bass attenuation of about 3 db. per octave in order to reduce the range of amplitudes covered by the significant components of speech. Experience has proved that for future work a still greater degree of attenuation would be desirable. Provision was also made in the recorder amplifier for volume compression, so that the weaker speech sounds could be recorded in sufficient strength to enable them to over-ride the thermal noise in the analyser circuits. This was adopted as a temporary expedient which should become unnecessary as soon as this noise level is sufficiently reduced. The speed V usually adopted for the recording film was 15 in/sec., so that a 3 kc/s. note, for example, was represented by a space periodicity of 0·005 in. The dispersion of the subsequent spectrum varies inversely as V, while there is a high-frequency cut-off to the recorded spectrum which varies directly as V; the selected value of 15 in/sec. was a convenient compromise for use in association with the optical system described in the next section.

(ii) *Analyser*: *Optical System*

The analysing process is most conveniently described in two sections, the first or optical stage being illustrated in Figure 1, the second or electrical stage being covered in the next section.

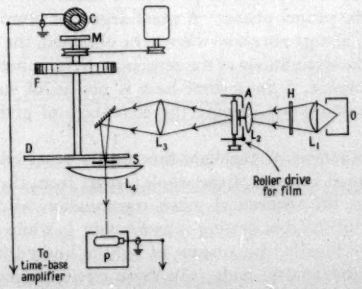


Figure 1. Optical system of analyser.

O is a strong source of light which, combined with the condenser lens L_1, gives a uniform illumination over that portion of the record film exposed by the aperture A; the width of the aperture is 5 mm., the same as that of the sound-track, while the length is of the order of 1 in., but this is adjustable. H is a filter of heat-absorbing glass, to protect the film. A continuous analysis is achieved by moving the film steadily past the aperture by means of a roller drive, which is actuated by a motor (not shown). A lens L_3 forms an image of the illuminated film, magnified approximately four diameters, on the rear surface of the disc D which is parallel to, and almost touching, the slit S. The deflecting mirror arrangement

became necessary when, in order to ensure constancy in the speed of rotation of the disc, a spring-loaded flywheel F was added to the disc shaft so forming an obstruction to any direct light. The field lens L_2 serves to concentrate the light on to the image-forming lens L_3, and behind the slit S is placed a condenser lens L_4 which finally converges the light on to the phototube P. This convergence takes place in a plane at right angles to the diagram, the source O being a line-source producing line-images at A and S. The disc D is a glass plate 6 in. in diameter; developed in a photographic emulsion on its lower surface there is a parallel-line grating having a sinusoidal variation of density (55 lines/inch). The full theory of the disc and its action is dealt with in § 3. Viewed through the narrow slit it presents the appearance of a sinusoidal grating whose periodicity varies from infinity to a lower limit of 1/55 in. as the disc rotates. The disc is motor-driven at a speed of 20 revolutions per second, and in every revolution the complete range of frequencies is explored four times, each quadrant of the disc being associated with one complete F-scan.

The changes produced in the phototube current as the scanning-frequency coincides with the various recorded frequencies in turn have to be amplified and combined with a suitable synchronized time-base to produce the desired F-spectrum on an oscillograph screen. The time-base signal needs to be synchronized and correctly phased with regard to the rotation of the disc; the constancy of the phase relation is of prime importance since any variations will lead to blurring of the final patterns. Such blurring occurred with earlier arrangements using separate oscillator circuits which were triggered by the rotation of the shaft, but ultimately the following arrangement proved both simple and satisfactory. A bar-magnet M attached to the shaft at right angles induces a voltage in the fixed pick-up coil C, which is adjusted in position to make the voltage as sinusoidal as possible and of the proper phase. A small degree of filtering in the time-base amplifier enables an almost pure sine wave to be obtained, the steadiness of whose phase is shown by the straightness of the zero lines in the subsequent FT-patterns. For the present purpose, a linear time-base is not at all suitable, a sinusoidal one derived in the above way having the advantage of giving a linear F-scale in the patterns.

The useful component of the light-flux finally received by the phototube represents only a small fraction of the whole; apart from the usual losses in the optical components, the theoretical mean transmission of the sound track is at best only $\frac{1}{2}$, that of the disc grating is again only $\frac{1}{2}$, while the practical values will be rather less. Finally, the amount of light is limited by the slit S, which is some 5 in. long and 1/60 in. wide. In these circumstances, it is a matter of no little difficulty to obtain a sufficient ratio of signal to thermal noise in the phototube and pre-amplifier circuits; in the present experiments the problem was accentuated by the fact that an electron-multiplier phototube could not be procured and conventional tubes had to be used. While it may be worth noting that the method can be made to work with ordinary phototubes, it is not a procedure to be recommended. It is necessary to use a light source of extremely high intensity, and one whose spectral characteristic is well matched to that of the photocell. The source finally used in this work was a Box Type M.E. 250 watt projection lamp, a high pressure mercury arc, which was operated on D.C., and used in conjunction with an RCA926 vacuum photocell having maximum sensitivity in the green. It was necessary to modify the apparatus to suit this particular

lamp, which only runs satisfactorily if the arc is vertical, so that the aperture A and the slit S needed to be vertical. The light distribution along the arc was not ideal, since there were fairly pronounced peaks of intensity towards the ends. These can be compensated for by using graded screens of the type mentioned in § 5 (ii).

(iii) *Analyser*: *Electrical System*

We now consider the way in which the phototube signals and the time-base signals are combined to produce the desired frequency-spectrum on the oscillograph screen; this is achieved by the arrangement shown in block form in Figure 2. The time-base signals from the pick-up coil, after amplifying and filtering, are applied to the X-plates of the oscillograph. Their sinusoidal character makes the to-and-fro motions of the spot occupy equal intervals of time, so that in effect two traces are superposed. To avoid this, the method finally adopted was to develop a square wave in time quadrature with the time-base wave, applying this to the Y-plates so as to deflect the spot off the screen during the return trace. As explained above, each quadrant of the disc performs a complete F-scan, so that although the period during which the spot is visible corresponds only to a

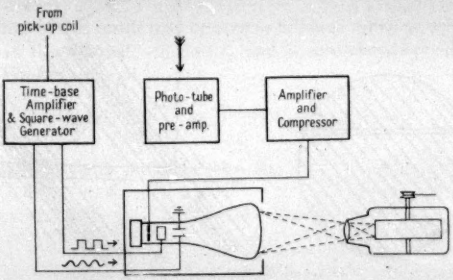


Figure 2. Electrical system of analyser.

half-rotation the frequency range is still explored twice during this period. It is convenient in theory to represent this double F-scan as a single scan from $F = -\infty$ to $F = +\infty$; the centre of the trace then corresponds to $F = 0$ and the regions on either side to negative and positive frequencies respectively. If the whole trace is photographed, one thus obtains a double record, this being the reason for the symmetrical nature of the FT-patterns reproduced in this paper.

The signals forming the pattern come from the phototube, which is mounted in a shielded box together with its pre-amplifier. The signals then go to the main amplifier and amplitude-compressor and finally to the grid of the cathode-ray oscillograph giving intensity-modulation of the trace. This is so arranged that when there are no signals the cathode-ray spot is completely suppressed. The amplitude-compressor was designed to meet the special requirements involved in recording intensity-modulated traces photographically. Some degree of compression is desirable, since a very large range of signal amplitudes is encountered, and if the weaker peaks are amplified sufficiently to permit of their being photographed, it is difficult to keep the cathode-ray tube focused for all the peaks. This compression is obtained by using a variable-mu tube with almost

zero bias working on the output from the negative peaks. The positive peaks are suppressed and damage to the tube from excessive grid-current avoided by using a grid-blocking resistor.

As the record film is moved through the light-beam by the roller drive, the space-function (originally a time-function) which is being analysed changes continuously, and the whole object of the method is to render the corresponding changes in the frequency-spectrum available for examination. The cathode-ray trace is therefore photographed by the recording camera shown in Figure 2, so building up a continuous FT-pattern. The camera drive is mechanically geared to the roller drive in Figure 1 so that the rates of movement of the sound film and the analysis film preserve a constant ratio. This ratio is adjustable, and determines whether the time-scale in the final record is relatively open or more compressed; it corresponds to what has been termed the aspect ratio in the Bell Telephone investigations.

§ 3. THE FREQUENCY-SCANNING DISC

(i) *Action and Theory*

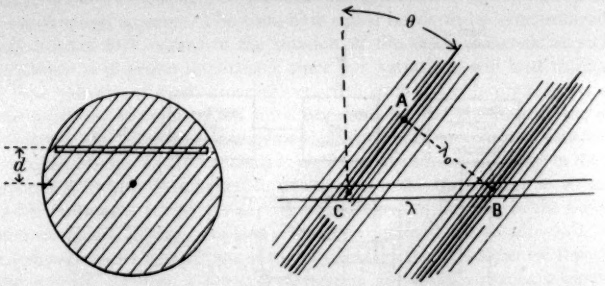


Figure 3.

λ_0 = space period of grating = AB. x = distance along slit from mid-point.
λ = space period along slit = BC. $-l_1$, l_2 = coordinates of ends of slit.
$\lambda = \lambda_0 \sec\theta$. $l = l_1 + l_2$ = length of slit.
d = perpendicular distance of slit from centre of disc.

The transmission functions of both disc and film may be described in terms of the single coordinate x, provided we discuss the film solely in terms of its magnified image at the slit. The film is moved relatively slowly, and to begin with we may regard it as stationary. The transmission functions (fraction of light transmitted) for the film record and the disc will be zero everywhere except between $-l_1$ and l_2. In this region let us write the transmission function for the film as $f(x)$ and that for the disc as $B[A+\cos(px-\epsilon_p)]$ where A and B are constants and

$$p = 2\pi/\lambda = p_0 \cos\theta \qquad \ldots\ldots(1)$$

where $p_0 = 2\pi/\lambda_0$. Thus p is a function of θ, varying harmonically between the limits $\pm 2\pi/\lambda_0$ as the disc rotates. The phase ϵ_p also varies during the rotation, although it can be made independent of θ if a diametral slit is used, as in the instrument of Born, Fürth and Pringle. While this is an advantage for computational work, it is unnecessary here, and the non-diametral slit is preferable

with the larger size of disc used by us since it enables a shaft drive to be employed. ϵ_p is therefore variable, and we merely note its value for reference purposes, namely

$$\epsilon_p = (2\pi d/\lambda_0)\sin\theta. \qquad \ldots\ldots(2)$$

We shall call this quantity ϵ_p the "instrumental phase".

The flux density at any point x on the slit is given, apart from constant factors, by the products of the transmissions of the film and disc,

$$L(x, p) = f(x)[A + \cos(px - \epsilon_p)].$$

The total light-flux through the slit at a particular instant is thus proportional to

$$L(p) = \int_{-l_1}^{l_2} L(x, p)\,dx = A\int_{-l_1}^{l_2} f(x)\,dx + \int_{-l_1}^{l_2} f(x)\cos(px - \epsilon_p)\,dx.$$

We disregard the first term, which having no p-factor is quasi-constant and will produce no amplified signal. The signal voltage applied to the cathode-ray oscillograph may thus be written, neglecting constants,

$$V(p) = \int_{-l_1}^{l_2} f(x)\cos(px - \epsilon_p)\,dx. \qquad \ldots\ldots(3)$$

The significance of this result may be seen as follows: when an arbitrary function $f(x)$ is given in the interval $-l_1 < x < l_2$, and is zero elsewhere, then a standard form for the Fourier integral is

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk \int_{-l_1}^{l_2} f(x')\cos k(x' - x)\,dx'$$

which on expansion yields the three relations

$$f(x) = \int_{-\infty}^{\infty} A(k)\cos(kx - \delta_k)\,dk, \qquad \ldots\ldots(4)$$

$$A(k) = \frac{1}{2\pi}\int_{-l_1}^{l_2} f(x')\cos(kx' - \delta_k)\,dx', \qquad \ldots\ldots(5)$$

$$\tan\delta_k = \int_{-l_1}^{l_2} f(x')\sin kx'\,dx' \Big/ \int_{-l_1}^{l_2} f(x')\cos kx'\,dx'. \qquad \ldots\ldots(6)$$

From a physical viewpoint the above equations are very appropriate to the present case. Thus if $f(x)$ represents the sound to be analysed, equation (4) expresses this as a spectrum ranging over all frequencies from $-\infty$ to $+\infty$, while (5) and (6) give the amplitude A and the phase constant δ of each component. (The quadrant in which δ lies is determined by the signs of the numerator and denominator in (6).) Now, the required amplitude A of each successive k-component will be given by the voltage applied to the cathode-ray oscillograph, since the rotation of the disc will cause the value of p in equation (3) to coincide with each k-value in turn, and then equations (3) and (5) are the same in all but one respect. The difference is in regard to the phase; when p coincides with k, the phase-constants ϵ_k and δ_k will usually not coincide. The latter is given by equation (6) and may be termed the "true phase" of the particular k-component; the former, as has been seen, depends upon the rather adventitious factor d, the distance of the slit from the axis of rotation, and is the "instrumental phase". The difference is of little significance however in the present method, for the following reasons

So far we have assumed the film-image to be stationary, but during an actual analysis it is moving continuously parallel to the slit and so across the rapidly spinning disc. Except for the lowest frequencies (below about 30 c/s.), any significant component will have a number of wavelengths present in the aperture, so that the motion of the film will cause the phase equality $\epsilon_k = \delta_k$ to be satisfied periodically. At such times, equations (3) and (5) completely coincide and the oscillograph will register the full amplitude A of the component in question. It may be shown that in general the quantity actually registered is $A' = A\cos(\epsilon_k - \delta_k)$, which varies harmonically between the limits A and $-A$; the negative signals, as previously explained, are not in fact utilized. In theory, therefore, each component is registered periodically, and if the film image is moved slowly enough the corresponding changes in spot brightness (or, using amplitude display, in line height) can easily be seen on the oscilloscope. If a spectrogram were made with a sufficiently open time scale, these fluctuations would impart a striated appearance to each trace, such that each striation corresponded to one period, on the time scale chosen, of the spectral component in question. This property may in later work be used to study phase relations in detail.

Figure 4. "You and I", spoken by male voice (Subject A). The almost constant pitch of each partial shows that the phrase was spoken with very little inflection. The first word is the diphthong ɪu. In the ɪ sound the first formant region is rather low and the second formant region is rather high, so that initially only the 2nd and 7th harmonics are visible. In the transition to the u sound the second formant descends so that the 7th, 6th and 5th harmonics are emphasized in turn. The lower formant simultaneously ascends, the 3rd and 4th harmonics being visible on the original record. At the beginning of the word "and", the harmonics from the 2nd to the 6th are visible, the 2nd and 4th being strongest; the transition to the nasal n sound is accompanied by a sudden fall in the first formant which makes the 3rd harmonic disappear abruptly in the dead space between the first and second formants. The word "I" is the diphthong aɪ. The a sound is compatible with three resonant regions, the 2nd, 4th and 7th harmonics being initially strongest; the subsequent transition to the ɪ sound of the diphthong is discussed below for Figure 5, in which the same diphthong appears with more detail.

Figure 5. "My poor boy" (Subject A), spoken with considerable inflection. Here each word contains a diphthong. In "my", the initial a sound shows three regions of maximum intensity situated around 600, 1,000, and 2,000 c/s. respectively. In the transition to the ɪ sound, the lowest region descends so that the fundamental tone, previously missing, finally becomes predominant; meanwhile the highest region ascends causing emphasis on the 9th, 10th and 11th harmonics in turn until the components become too weak to register. The middle region appears to ascend towards the position previously occupied by the highest region, causing a momentary reappearance of a harmonic around 2,000 c/s. In "poor" the two sounds comprised in the diphthong uə are not strongly differentiated; as the pitch inflection rises, the 2nd and 3rd harmonics continue strong, but the 4th and 5th harmonics have to cross a dead space before they again reach a favourable region around 2,000 c/s. giving them a rather broken appearance. In "boy" the diphthong ɔɪ is marked by more obvious movements of the formant regions; two regions, initially nearly coincident, bifurcate to form distinct branches, one descending and one ascending, and there are broken traces of higher regions. A further note on this pattern appears in § 5.

Some objection might at first be raised to a procedure which, if the film is stationary, does not register all the components in the corresponding $f(x)$ simultaneously. It might be argued that as soon as the film begins to move, the function which fills the aperture is no longer quite the same $f(x)$, so that at no time will the analysis be the "absolutely correct" one. From this viewpoint, the worst cases would be those where there is a sharp pulse or discontinuity

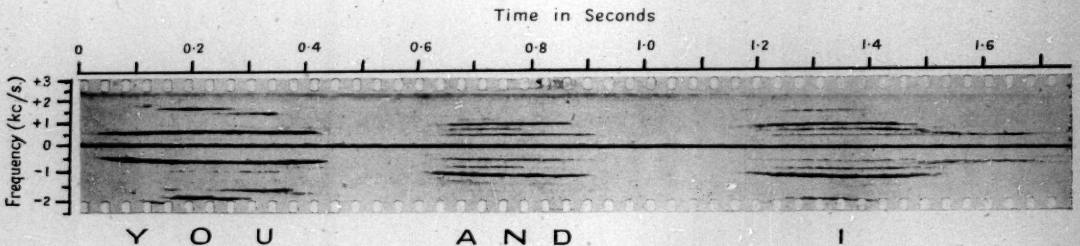


Figure 4. Frequency-time (FT) record for the spoken words "You and I". The instrument produces a double trace, indicated on the record by positive and negative frequency scales

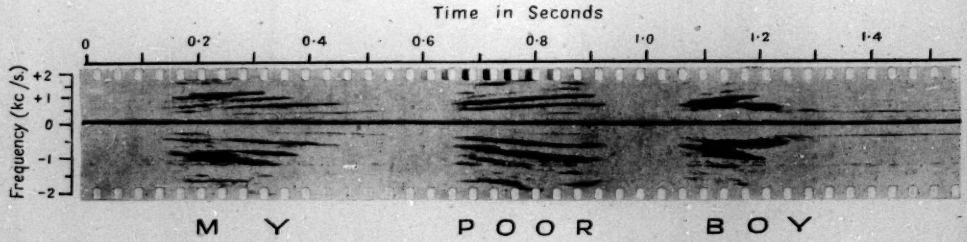


Figure 5. FT-record for the words "My poor Boy", each of which contains a diphthong.

To face page 156

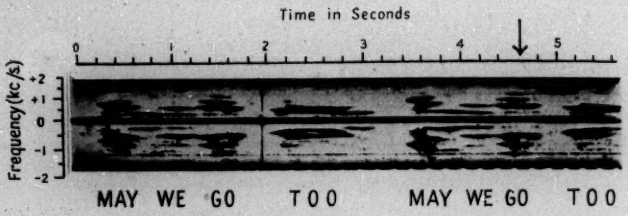


Figure 6. FT-record for the words "May we go too?", repeated. The arrow shows the location corresponding to the typical amplitude-spectrum shown in Figure 7 (*b*).

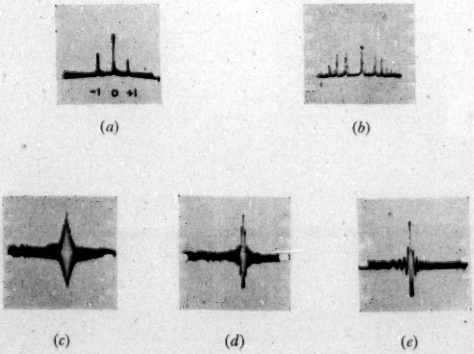


Figure 7. (*a*) Analysis of 1 kc/s. test-note. (*b*) Spectrum of portion of the word "go" in Figure 6, showing amplitudes of components. (*c*), (*d*), (*e*) Expanded profiles of spectral lines.

which a slight movement of the film might cause to move on or off the aperture. But the higher-frequency components, which will be the principal ones associated with such a pulse, will all achieve phase-coincidence many times as the pulse traverses the length l, and only those whose space period exceeds $l/2$ may fail to achieve coincidence. To measure these lowest frequencies accurately, one would naturally need to make l large; but this is merely an example of the acoustic uncertainty principle which affects all types of analyser.

With regard to the time-base, since in equation (1) we have $p=p_0 \cos\theta$, the procedure is merely to adjust the phasing so that the deflection X of the cathode ray spot is proportional to $\cos\theta$; the relation between X and p being then linear, the resulting spectra have a uniform frequency scale.

(ii) *Construction*

The discs so far used have been made as follows. A master-plate was prepared by a commercial photo-engraver, in the form of a grid with alternately opaque and transparent lines of equal width. Diffuse photographic copies were made of this with various conditions of lighting, and eventually a number of plates were obtained which appeared closely sinusoidal. Discs cut from these were tested by carrying out trial analyses of a commercial sound film of the type used for testing purposes, representing a pure 1,000 c/s. tone. To make the tests

Figure 6. "May we go too?", repeated, male voice (Subject B). This photograph is given to illustrate the following further points: (i) the repetition of the same words provides material for the recognition and association of patterns; (ii) the adoption of a different speed-ratio leads in this case to a more compressed time-scale—the greater compactness of the traces appears to aid in the recognition of patterns; (iii) measurements of the amplitudes of the various components are possible at any stage, in the form of "still" photographs obtained by switching over to amplitude-modulation on the cathode-ray oscillograph. As an example, a quantitative cross-section through the word "go" in this figure, at the point indicated by the arrow, is given in Figure 7(*b*). Keeping to the subject of speech sounds, such measurements would be of value in studying the frequency-response, and hence the damping-time, of the cavity-resonators. A calibration process would be necessary in conjunction with such studies in view of the bass-attenuation used in making the sound films, and of the amplitude compression in the signal amplifiers.

quantitative, the connections to the oscillograph were modified so as to produce amplitude-modulation of the trace on the screen instead of intensity-modulation. Figure 7 (*a*) is a typical trace, which in addition to showing the satisfactory nature of the disc used in these experiments, also shows another characteristic feature which appears in all the spectra. Any function $f(x)$ which is recorded appears as a certain modulation of the light transmission of the sound-film, and this, being essentially positive, must contain a large constant term. There is thus a strong steady signal corresponding to zero frequency in all the records, which incidentally is of value as a reference mark. Strong signals corresponding to the 1,000 c/s. test-note are visible on either side of this zero peak, without any extra or spurious signals, showing the disc-pattern to be very closely sinusoidal.

§4. SPECIMEN ANALYSES OF ACOUSTIC WAVEFORMS: SPEECH PATTERNS

Figures 4 to 6 show some typical results obtained with the one experimenta machine we have so far constructed; they are the FT-diagrams produced by short phrases of speech, spoken fairly deliberately, so that the patterns formed

by the individual words are for the most part well separated. The traces convey even on casual inspection much more information than an ordinary wave-form rendering of the sounds, and the information is more accurate than that sometimes obtained with a great deal of labour by using harmonic analysers to investigate successive portions of the wave-form.

The work of numerous previous investigators on the mechanism of speech leads one to look for certain major characteristics in the spectra first of all. In normal speech a stream of air is forced through the glottis, the narrow slit which separates the vocal cords, causing these cords to vibrate and setting up sound waves. The quality of the sound will depend on the mode of vibration of the vocal cords, and to a still greater extent upon the modifying action of the cavities formed by the tongue, lips etc. The fundamental or pitch note emitted by the vocal cords is accompanied by a large number of upper partials, and it has been established that there are certain frequency regions in which the amplitude of the partials is enhanced by resonance effects; these are known as formant regions, since they determine the quality of the particular sound. The formant frequencies for the spoken vowels have been estimated by various workers, the conclusions of Paget, Crandall, Stewart and Steinberg, for example, being summarized in a paper by Steinberg (1934), and more recent results are given by Potter *et al.* (1946). The estimates differ somewhat, but there are broad lines of agreement between them, which perhaps is all one can expect since, for example, the formant regions are affected by the nature of the preceding and following sounds. We may conclude this brief summary by a reference to the consonant sounds such as *p*, *t*, *k*, in which the air stream is interrupted or released very abruptly. A detailed study of consonants will demand a high resolution in the T-direction, whose adoption, however, will automatically result in a low frequency-resolution throughout the record owing to the uncertainty principle. In the photographs reproduced here the uncertainty in the T-coordinate is about 1/15 sec.; this is too large for consonants, but was chosen as a convenient mean value for preliminary studies in the spectrum as a whole. The photographs represent the voices of only two speakers, so that the spectra may not be entirely typical, but they serve to illustrate the potentialities of the method and the general agreement with earlier findings where these are comparable.

§ 5. THE FORMATION OF SPECTRAL LINES

(i) *Resolving Power of Instrument*

In optics, when light of constant frequency is examined with a spectroscope, the best approximation to a spectral line is usually a $(\sin u/u)^2$ intensity-function, the central peak of which represents the position of the ideal "line".

In the present acoustic spectrograph, the corresponding case would be the analysis of a film containing a pure sine or cosine wave, which apart from unimportant constants may be represented by $C+\cos(kx-\delta_k)$, the transmission of the disc being represented as before by $A+\cos(px-\epsilon_p)$. The light to the photocell is then given by

$$L(p,k)=\int_{-l_1}^{l_2}[C+\cos(kx-\delta_k)][A+\cos(px-\epsilon_p)]\,dx$$

$$=L_0+L_1(p)+L_2(k)+L_3(p+k)+L_4(p-k),$$

since on integration the expression resolves itself into a number of terms typified by L_4 which when written in full gives:

$$L_4(p-k)=(p-k)^{-1}\sin[(p-k)(l_2+l_1)/2]\cos[(p-k)(l_2-l_1)/2-\epsilon_p+\delta_k]. \quad \ldots\ldots(7)$$

We will confine our attention to this term, since it is the only one to undergo large variations when the scanning-frequency p approaches the particular component-frequency k; the other terms are constant or quasi-constant for small changes in p. Of course, when we consider the negative values of p, which correspond in the photographs to the traces below the zero line, the rôles played by the terms $L_3(p+k)$ and $L_4(p-k)$ become interchanged, but this does not affect the argument. By analogy with optical spectra, the variation of L_4 with p may be called the "line profile" of the k-component line. As we are only interested in the behaviour in the immediate neighbourhood of the line, we may transfer our frequency-origin to this point and use instead of p a variable u defined by $u=(p-k)l/2$. We must consider the relative rates of variation with u of the sine and cosine factors in equation (7), with due regard to the fact that in practice $l_1\simeq l_2$. Thus the sine factor varies only with u, while the variation in the cosine factor arises mainly from ϵ_p. Investigation shows that there are three rough categories for the response of the system to any particular component k, depending on the value of the ratio k/p_0 where p_0 is the maximum value attained by p in the scanning process. There is a critical value which arises from the use of the non-diametral slit, namely: $k/p_0=2d(l^2+4d^2)^{-\frac{1}{2}}=\alpha$, say.

Case 1. Low frequency region. $(k/p_0\ll\alpha)$. The variation of the sine term predominates, causing the line profile to follow a $\sin u/u$ function as in optics.

Case 2. Intermediate region. $(k/p_0\simeq\alpha)$. Both sine and cosine terms vary at comparable rates.

Case 3. High frequency region. $(k/p_0\gg\alpha)$. The cosine term varies much more rapidly than the sine term, the latter merely giving the envelope of the line profile which is a $\sin u/u$ function as before (Case 1).

It has been possible to test these conclusions to some extent by deliberately reducing the resolving power (length of aperture A reduced to 0·05 in.), thereby broadening the spectral lines sufficiently to reveal their actual profiles. The traces in Figure 7 (*c*), (*d*) and (*e*), taken with amplitude modulation, illustrate the responses to high, moderately high, and intermediate acoustic frequencies respectively, and show the decrease in the rate of oscillation of the cosine term with decreasing acoustic frequency. The experimental arrangement gave $\alpha\simeq 0{\cdot}6$, so the transition from a "$\sin u/u$-profile" to a "$\sin u/u$-envelope" occurred at about 60% of the maximum frequency. It is clear that we can regard the $\sin u/u$ function as determining the effective line-breadth for all frequencies, since any finer structure within the envelope will not affect our ability to resolve lines. We adopt the usual convention that two frequencies differing by ΔF are considered separate if the peak of the pattern for one coincides with the first minimum of the other, which corresponds to a range of π in u. We thus obtain

$$u=(p-k)l/2=2\pi(\Delta F/mV)(l/2)=\pi, \quad \text{or} \quad \Delta F=mV/l \qquad \ldots\ldots(8)$$

where V is the film-velocity used during recording, l is the length of the slit, and m is the linear magnification. Hence we get the result: *the frequency-resolution of the instrument is constant over the whole frequency range.*

The above result is really an expression of the uncertainty principle which links the F-resolution and the T-resolution of the instrument. The length l' of the sound film within the aperture will correspond to a certain interval of time given by $\Delta T = l'/V$. Hence from (8), and with our definitions of ΔF and ΔT, we have $\Delta F \Delta T = 1$. This expression determines the theoretical value of the line-breadth, without reference to instrumental factors such as the size of the cathode-ray spot which may tend to increase the actual breadth.

Brief reference should be made to an aspect of the uncertainty principle which is illustrated in certain of the photographs reproduced in this paper, namely an occasional blurring of the traces which becomes most marked whenever the pitch is changing rapidly. Some time before the introduction of practical methods for making FT-diagrams, Gabor (1946) gave theoretical diagrams for a number of typical phenomena, and Figure 1.12 in his paper shows a diagram for a frequency-modulated signal. In the regions where the "frequency" is changing most rapidly an inevitable blurring occurs, which can never be eliminated, although it can be minimized by a suitable choice of the $\Delta T/\Delta F$ ratio. It is interesting to compare the figure above cited, with the practical example afforded by the word "boy" in Figure 5, in which the pitch first rises and then falls again in about 1/5 sec. In such cases any attempt to approximate to the frequency at a particular instant by reducing ΔT will ultimately be defeated by the associated increase in ΔF.

(ii) *Resolving Power with Special Apertures*

A potentially valuable property of the present analyser is the ease with which the whole nature of the T-response and the associated F-response can be controlled. Instead of the simple rectangular aperture A we may substitute either an aperture whose height h varies along its length according to any desired function of x, or an aperture with a screen of graded density which varies in the desired way with x. The disadvantage of a simple rectangular aperture is that it produces an F-response curve which is comparatively complex, as discussed in the last section. Since we are interested in the FT-diagram it would be more desirable to have symmetry between the F and T response curves, which from the theory of the instrument comprise a pair of Fourier transforms. The self-reciprocal character of the error-function in this transformation makes it natural to consider its adoption for the shape of the response curves.

An immediate advantage of this arrangement over the unsymmetrical one previously used, would be the more rapid falling-off of the line-profile curves; instead of the factor $\sin u/u$ we should have $\exp(-\alpha u^2)$. This would be of benefit in distinguishing small signal peaks in the proximity of large peaks, since the subsidiary maxima of the latter, which might obscure a small signal, would be eliminated. Gabor (1946) in his very fundamental investigation of the theory of FT-diagrams, has in fact shown the error-curve to be the function satisfying the condition for minimum uncertainty, and leading to minimum areas in the diagram.

While it is not actually possible to employ infinite error-functions, approximations of the form $\cos^n \pi x/l$ should give satisfactory results with $n = 2$ or 3, the curves being terminated at $x = \pm l/2$. Examples of such curves are shown

in Figure 8(*a*), fitted into the original rectangular aperture of length l, together with an error curve falling to 1% of its maximum value in the same range. In Figure 8(*b*) are shown the corresponding line profiles which fall off as $u^{-(n+1)}$. It will be seen that there is little to choose between the graph for $n=3$ and the infinite error function. It is hoped to include this type of provision in future experiments. The value of l could be chosen to give any desired time-resolution, and altogether the arrangement would be much simpler than with electrical analysers where it would be difficult to alter the damping times or the character of the response curves in the same way.

(iii) *Amplifier Bandwidth Requirements*

The amount of detail required in the spectrum, and the speed with which it is desired to carry out the analysis, both affect the amount of information the photocell amplifier is called upon to transmit in a given time, and therefore the bandwidth required. In order to estimate the probable bandwidth, we could proceed to express the signal imparted to the cathode-ray oscillograph given by equation (3) as a function of time, but this method is not very convenient as it leads to rather complicated expressions involving Bessel functions. A simpler method of approach is to assume $f(x)$ to consist of discrete sinusoidal components, and find the form of the amplifier-pulse when the search-frequency p traverses

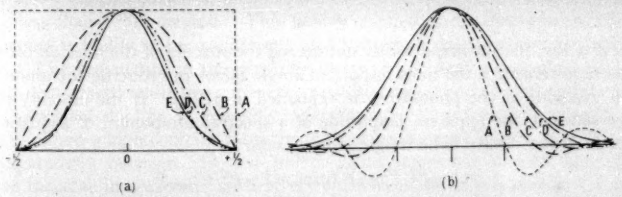


Figure 8.

(*a*) Aperture profiles. Curves A, B, C, D are profiles of form $\cos^n \pi x/l$ for $n=0, 1, 2, 3$ respectively; Curve E is error-curve.

(*b*) Corresponding spectral-line profiles.

a particular frequency k, by using equation (7). If the line profile of § 5, Case 1, is assumed, an approximation can be used for p in the neighbourhood of k so that the pulse has the same analytical form expressed as an approximate function of time as it had as a function of θ. From the reciprocal property of Fourier transforms it can then be shown that the frequency-band extends uniformly from zero to a sharp cut-off at a frequency f_1 given by $f_1=(lp_0\omega \sin\theta_k)/4\pi$ where ω is the angular velocity of the disc. This has its maximum value when $\theta=\pi/2$ corresponding to $k=0$ and here the assumptions made are well justified.

At the other extreme represented by § 5, Case 3, we have $\theta=0$ and the pulse is dominated by the cosine term through the instrumental phase ϵ_0. The resulting frequency is that with which the disc lines sweep across the slit, namely $f_2=dp_0\omega/2\pi$. For our instrument, f_1 and f_2 worked out at about 17 and 9 kc/s. respectively, with the disc rotating at 20 rev/sec. The maximum frequency of 17 kc/s. is readily handled by carefully designed resistance-capacitance coupled amplifiers.

(iv) *Limitation on Slit-Breadth arising from* ϵ_p

It has been shown in § 3 that as a result of the formal resemblance between equations (3) and (5), the analyser theoretically gives not only the frequencies but also the amplitudes of the various components present in $f(x)$. In normal FT-records the amplitudes are registered by the intensity of the oscillograph spot, although they may also be measured quantitatively if desired, as in the example in Figure 7 (*b*). In deriving equation (3) however, we have assumed the slit S to be of negligible breadth, in which case ϵ_p in equations (2) and (3) may be regarded as constant over the breadth of the slit. We now consider the general case of a slit of finite breadth b, and determine its light transmission as a function of x and p. The contribution of an elementary strip of breadth dr, whose distance from the centre of the disc is r, may be written $K_p \cos(px-\epsilon'_p)\,dr$ where $\epsilon'_p=(2\pi r \sin\theta)/\lambda_0$ and p and K_p are independent of r. The contribution of the whole slit will be $\int_{d-b/2}^{d+b/2} K_p \cos(px-\epsilon'_p)\,dr$. On evaluating this it is found that the whole slit gives rise to a transmission function similar to that of an elementary strip, with a phase equal to that of the mid-strip, and an amplitude given by $[K_p b \sin\{(\pi b \sin\theta)/\lambda_0\}][(/\pi b \sin\theta)/\lambda_0] = K_p b_{\text{eff}}$, where b_{eff} is the effective slit breadth. Since this cannot be allowed to become zero, we must have $(\pi b \sin\theta)/\lambda_0 \leq \pi$ or since the maximum value of $\sin\theta$ is unity

$$b \leq \lambda_0, \qquad \ldots\ldots(9)$$

that is to say, the slit-breadth may not exceed the spacing of the lines on the disc. This requirement is the most important single factor in reducing the amount of light available at the phototube, as explained in § 2 (ii). If the equality in (9) were satisfied, the apparent amplitude of a spectral component of acoustic frequency F would have to be weighted by a factor

$$\pi(1-F^2/F^2_{\max})^{\frac{1}{2}} \operatorname{cosec} \pi(1-F^2/F^2_{\max})^{\frac{1}{2}}$$

where $F_{\max}$ is the maximum frequency the instrument can detect. In effect, the use of the wider slit brings about a not altogether undesirable bass attenuation, which is additional to such attenuation as may be introduced deliberately during the recording process. Should quantitative amplitude measurements be desired, the calibration mentioned in § 4 would have to be supplemented so as to include this slit-breadth factor. A calibration including both factors could easily be obtained experimentally by recording and analysing a wave-form such as a repeated square-wave, whose exact spectral composition is known from theory.

REFERENCES

Brown, D., 1939, *Proc. Phys. Soc.*, **51**, 244.

Brown, D., and Lyttleton, J. W., 1947, *Nature, Lond.*, **160**, 709.

Fürth, R., and Pringle, R. W., 1946, *Phil. Mag.*, **37**, 1.

Gabor, D., 1946, *J. Instn. Elect. Engrs.*, **93**, 429.

Potter, R. K., *et al.*, 1946, *J. Acoust. Soc. Amer.*, **18**, 1 (series of six papers by Bell Telephone Laboratories).

Schouten, J. F., 1939, *Philips Tech. Rev.*, **4**, 290.

Steinberg, J. C., 1934, *J. Acoust. Soc. Amer.*, **6**, 16.

The Cause of Anisotropy in Permanent Magnet Alloys

By K. HOSELITZ and M. McCAIG
Permanent Magnet Association, Central Research Laboratory, Sheffield

MS. received 24*th June* 1948, *in amended form* 9*th September* 1948

ABSTRACT. Measurements of magnetostriction have been made on samples of an anisotropic permanent magnet alloy. From the results it is concluded that the domain magnetization is, in the absence of a field, along that easy crystallographic direction which makes the smallest angle with the axis of anisotropy. This view is supported by measurements of the remanence. The ideal arrangement which obtains after hardening seems to be slightly disturbed during the subsequent tempering operation.

§ 1. INTRODUCTION

THE most powerful permanent magnets now made commercially are subjected to the following processes in order to obtain the required properties (Hoselitz 1946). The magnet is heated to a high temperature, usually about 1,300° c., and cooled at a controlled rate while situated in a magnetic field of not less than 1,000 oersted. This field is applied in the direction in which the magnet is finally to be magnetized, and as a result the alloy becomes anisotropic. The remanence in the direction in which the field was applied during cooling is increased while the remanence in a direction at right angles is reduced. This controlled cooling is generally referred to as "hardening". Finally in order to obtain a high coercive force the magnet is tempered for a suitable period at temperatures between 550 and 700° c. This tempering process gives the desired increase in coercivity with a very small decrease in remanence. There are several new commercial alloys which are produced in the above manner, although the precise details of composition and heat treatment vary slightly.

In order to obtain information about the nature of the processes producing the anisotropic properties, magnetostriction measurements have been made on samples of the alloy known as 'Alcomax II'.

§ 2. EXPERIMENTAL PROCEDURE

The composition of the alloy was Ni 11·71%, Co 22·09%, Cu 4·59%, Al 7·96%, Si 0·17%, Ti 0·42%, balance Fe.

The samples were in the form of rectangular test pieces of approximate dimensions $4 \times 3{\cdot}5 \times 1{\cdot}0$ cm. We shall refer to directions parallel to these lengths as OX, OY, OZ respectively. An electromagnet with cylindrical pole-pieces 10 cm. in diameter with a gap of about 7 cm. was used to supply the field. The specimen was clamped at one end to a firm support so that it was held in a horizontal position between the polefaces, the other end being left free. The change in length during magnetization was measured by a modification of Honda's optical method. A bridge consisting of a brass bar of square cross section was placed on the specimen. One end of this bar was fitted with a sharp V-shaped edge which rested on the specimen, the weight of the bar being sufficient to prevent this edge from slipping on the specimen. The other end of the bar was

carried on a roller 0·81 mm. in diameter; this roller also rested on the specimen and in consequence rotated as the specimen expanded or contracted. Thus the bridge could be placed in any position relative to the field direction, and transverse as well as longitudinal magnetostriction could be measured with the same apparatus.

The angle of rotation of the roller, which is a measure of the change in length of the specimen, was measured as follows: A small concave galvanometer mirror was fixed to the end of the roller with its plane parallel to the axis of the roller. Light from a galvanometer lamp was shone on this mirror and the reflected image was observed by a travelling microscope capable of reading a displacement of 0·01 mm. For the measurement of the transverse magnetostriction the optical path was thus conveniently perpendicular to the axis of the polepieces. But when the longitudinal effect was investigated both the incident and reflected beams of light had to be deflected through a right angle by means of a small additional plane mirror placed close to the galvanometer mirror with its plane at 45° to the axis of the magnetic field. The length of the optical lever was a little more than one metre, and a magnetostriction of 10^{-5} gave a movement of the image of about 0·7 mm. No very great accuracy is claimed for these experiments, but the changes in magnetostriction to be reported are so large that an accuracy of say 10% is quite adequate for our purpose. Readings were repeatable to within ± 5 scale divisions. There was no noticeable effect due to heating up of the magnet but experiments could only be performed when the room temperature was reasonably steady.

Magnetostriction measurements were carried out for fields from 0 to 4,500 oersted, the field being measured directly by a magnetic potentiometer as described by Margerison and Sucksmith (1946). Both the longitudinal and transverse magnetostriction were investigated in the following way. The sample was demagnetized *in situ* by repeatedly reversing the applied field, beginning with high fields sufficient to produce saturation and decreasing this in small steps to zero. Demagnetization in this manner was found to be not quite complete in all cases, but was probably sufficient for the purpose of these experiments, since hysteresis effects were not more than 10% of the saturation magnetostriction and no magnetostriction could be detected for fields below 200 oersted. The curves showing magnetostriction against field were obtained by applying a progressively increasing field to a specimen, previously demagnetized in the way described above.

§ 3. RESULTS

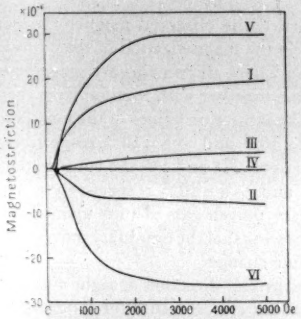
In Figure 1, curves I and II give the longitudinal and transverse magnetostriction for an isotropic sample either as cast or hardened without a magnetic field. Since the sample is isotropic similar results are obtained if the field is applied in the OY direction except of course that curve I is now obtained for the OY and curve II for the OX direction. The fact that saturation is effectively reached is evidence that a true magnetostriction is being measured and that there is no serious error due to the elastic extension of the sample subjected to magnetic forces.

Experiments were then made on a sample after cooling in a magnetic field, but before tempering. The results are shown in curves III, IV, V and VI of Figure 1. In these experiments the preferred direction, i.e. the direction of the

field applied during cooling, was along OX. The following significant changes compared with the isotropic case are evident. When the field is applied in the preferred direction the longitudinal magnetostriction is small and the transverse magnetostriction is zero, but when the field is applied in the OY direction, i.e. perpendicular to the preferred direction, the longitudinal magnetostriction is 50% greater than for the isotropic test piece and the transverse magnetostriction is only slightly less than the longitudinal.

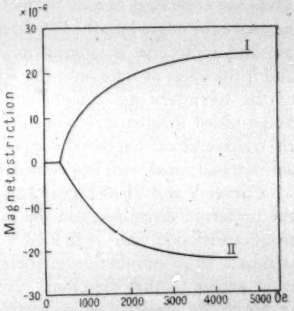
§4. DISCUSSION

The structure of Alcomax is known from x-ray work by Oliver and Goldschmidt (1946) to be body-centred cubic, and it is reasonable to assume that the easy direction of magnetization is along a cube edge as with iron. (Actually it would make very little difference from the point of view of this theory if we



Curve	Preferred Direction	Direction of Field	Direction of Measurement
I	Nil	0X	0X
II	Nil	0X	0Y
III	0X	0X	0X
IV	0X	0X	0Y
V	0X	0Y	0Y
VI	0X	0Y	0X

Figure 1.



Curve	Preferred Direction	Direction of Field	Direction of Measurement
I	0X	0Y	0Y
II	0X	0Y	0X

Figure 2.

assumed the easy direction to be a cube diagonal, as this would only make four instead of three easy directions. The theory would not, however, predict anisotropic properties for a hexagonal crystal such as cobalt, which has only one easy direction of magnetization). We know that the sample is polycrystalline and it is reasonable to assume that the orientation of the crystals is random. This will be true even after cooling in a magnetic field, since repeated experiments have shown that no recrystallization takes place during this treatment. In an isotropic sample there may, in a given crystal, be one cube edge along which the magnetization of the domains is most likely to lie. But over the whole sample the orientation of these directions of domain orientation is random.

We will show below that the experimental results can be satisfactorily explained if we assume that in the sample which has been cooled in a magnetic

field the domains in each crystal are magnetized along that cube edge which is in nearest alignment with the direction in which the field was applied during cooling.

In the domain theory of ferromagnetism three elementary processes are assumed—180° boundary movements, 90° boundary movements and rotation of the direction of magnetization from an easy direction to that of the applied field. Of these processes only the last two can produce magnetostriction. Consider the process which gave rise to curves III and IV of Figure 1. Since we have assumed that the domains are already parallel to the cube edges nearest to the OX direction no 90° changes can occur. The observed magnetostriction is therefore due solely to rotation of magnetization from the nearest cube edge into the OX direction. The saturation values are 3×10^{-6} for the longitudinal effect and zero for the transverse magnetostriction. These results indicate a small volume magnetostriction.

Turning back to curves I and II of Figure 1, the observed magnetostriction is due to some 90° boundary movements taking the magnetization of the domains from a random* distribution to a distribution where all domains are magnetized along the cube edge nearest to the OX direction together with rotation from these nearest cube edges into the OX direction. The saturation values of the longitudinal and transverse magnetostriction are 20×10^{-6} and -8×10^{-6} respectively, and if the effect of rotation is the same as for curves III and IV, the contributions of the boundary movements are 17×10^{-6} and -8×10^{-6} respectively. The longitudinal magnetostriction due to boundary movements is thus about twice the transverse value and of the opposite sign, showing that the boundary movements are not associated with any considerable volume change.

Curves V and VI of Figure 1 are for H in the OY direction at right angles to the preferred direction, and the saturation values are 30×10^{-6} for longitudinal magnetostriction and -26×10^{-6} for the transverse magnetostriction. There is thus a large contraction in the preferred direction when a field is applied at right angles to that direction. Now the magnetostriction in this arrangement is due to the direction of magnetization in the domains changing from the cube edge nearest to the OX direction to actual parallelism with the OY direction.

These changes in length can be regarded as equal to the sum of those due to taking the domains from the cube edge nearest to the OX direction into a random orientation plus those due to the changes from random orientation to parallelism with the OY direction.

If we express these relationships algebraically we have λ_{I}, λ_{II}, λ_{III}, λ_{IV}, λ_{V} and λ_{VI} to represent the saturation magnetostriction obtained from curves I, II, III, IV, V and VI respectively:

$$\lambda_{\mathrm{V}} = \lambda_{\mathrm{I}} - \lambda_{\mathrm{II}} + \lambda_{\mathrm{IV}} = (20 + 8 + 0) \times 10^{-6} = 28 \times 10^{-6},$$

compared with the measured value of 30×10^{-6}, and

$$\lambda_{\mathrm{VI}} = \lambda_{\mathrm{II}} - \lambda_{\mathrm{I}} + \lambda_{\mathrm{III}} = (-8 - 20 + 3) \times 10^{-6} = -25 \times 10^{-6},$$

compared with the measured value of -26×10^{-6}. Thus our hypothesis about domain magnetization explains the changes in magnetostriction due to heat treatment in a magnetic field. It is significant that when the field is applied at

* It is possible that the distribution of the domains in the demagnetized state was not completely random, but contained an excess in directions parallel to those of the alternating demagnetizing field.

right angles to the preferred direction, the transverse magnetostriction is only a little less than the longitudinal. This result does not imply that there is a large volume magnetostriction but rather that all the domains were originally magnetized substantially in the OX direction and the field has turned the magnetization into the OY direction leaving the OZ direction unaffected. If the magnetostriction had been measured in the OZ direction it would have been very small. This has been confirmed by experiment on a sample in which OZ was made the preferred direction during hardening in a magnetic field. H was applied along OX and the saturation magnetostriction along OY was only $2{\cdot}5 \times 10^{-6}$.

The suggested explanation of the effect of the magnetic field applied during hardening on the distribution of domain magnetization is supported by considering the ratio J_r/J_s, the intensity of magnetization at remanence to that at saturation. Assuming that the domains are all parallel to the cube edge nearest to the OX direction, J_r/J_s can be calculated to be 0·831 when magnetized in the OY or OZ directions. (For details of the calculation see the appendix.) The actual values of J_r/J_s for the specimen used after cooling in a magnetic field were: OX direction, 0·87; OY direction, 0·27. For an isotropic specimen cooled without a field the value is 0·67.

If all 90° boundary movements are completely reversible, the value of J_r/J_s should be 0·5. However, it is not surprising that this is not the case in the majority of magnetically hard substances, and in fact measured values for permanent magnet materials show that J_r/J_s lies mostly between 0·6 and 0·7. Thus the value given here for isotropic Alcomax II does not indicate any difference in character between this alloy and those that are not capable of being made anisotropic.

§ 5. EFFECT OF TEMPERING

During the tempering treatment, which has the main object of increasing the coercive force, certain other changes take place. Thus the values of J_r/J_s become: OX direction, 0·81; OY direction, 0·38; isotropic specimen, 0·55.

The results suggest that some of the anisotropic properties are lost during tempering and particularly that when a field has been applied in the OY or OZ direction magnetization in all the domains does not immediately return to the cube edge nearest to the preferred OX direction.

Magnetostriction experiments were also carried out on a tempered specimen with the preferred direction again along OX and the results are shown in Figure 2.

When the field during measurements is in the preferred direction the results are not very different from those obtained before tempering, but when the field is in the OY direction the magnetostriction is less than before tempering. The transverse magnetostriction is nearly equal but opposite in sign to the longitudinal. It is not easy to account precisely for all these changes, but it does appear that the ideal arrangement of the domains postulated after hardening is disturbed to some extent by tempering.

The practical effect of cooling in a magnetic field is to increase the remanence by aligning the domains along the easy directions nearest to the preferred direction. The treatment does not of itself greatly increase the coercive force, but because it discourages 90° boundary movements, the squareness of the demagnetization curve is increased and this greatly improves the value of $(BH)_{max}$. Most of this advantage is retained during tempering, the purpose of which is the increase in the coercive force.

We have not in this paper attempted to distinguish whether this partial alignment of the domains in the anisotropic magnet is due to crystalline energy, magnetostrictive stresses, or shape anisotropy of single domain particles (Stoner and Wohlfarth 1948), but have been content to point out the nature of the alignment which any theory must seek to explain.

ACKNOWLEDGMENT

The authors thank the Permanent Magnet Association for permission to publish this paper.

REFERENCES

HOSELITZ, K., 1946, *J. Sci. Instrum.*, **23**, 65.
MARGERISON, T. A., and SUCKSMITH, W., 1946, *J. Sci. Instrum.*, **23**, 182.
OLIVER, D. A., and GOLDSCHMIDT, H. J., 1946, *E.R.A. Report N/T.* 41.
STONER, E. C., and WOHLFARTH, E. P., 1948, *Phil. Trans. Roy. Soc.*, **240**, 599.

APPENDIX

Let OX be the preferred direction of magnetization. Let there be N equal elements of volume ΔV per cm^3 over which the magnetization is uniform. These elements may be equal to or smaller than the domains. Let J_s be the saturation intensity of magnetization and J_r the intensity of magnetization at the remanence point. Assume that the alloy is polycrystalline, with a cubic lattice, and that the orientation of the crystals is random.

Case 1. The alloy has been magnetized in the direction OX and is at the remanence point. If each element of volume is magnetized parallel to a cube edge there are $3N$ possible directions per cm^3 with a positive component along OX. In any crystal let α_1 be the cosine of the smallest angle between a cube edge and OX and let α_2, α_3 be the cosines of the angles between the other cube edges and OX. In an anisotropic alloy

$$J_r = \Sigma J_s \alpha_1 \Delta V, \qquad \ldots\ldots(1)$$

summed over unit volume. The number of cube edges making an angle between θ and $\theta + d\theta$ with OX is $3N \sin\theta\, d\theta$. Out of these cube edges let the number which are the cube edges nearest to OX be

$$3N \sin\theta f(\alpha_1)\, d\theta = -3Nf(\alpha_1)\, d\alpha_1. \qquad \ldots\ldots(2)$$

$f(\alpha_1)$ is the probability that a cube edge making an angle $\cos^{-1}\alpha_1$ with OX is a nearest cube edge to OX, that is, the probability that α_1 is greater than α_2 or α_3. Now

$$\alpha_1^2 + \alpha_2^2 + \alpha_3^2 = 1. \qquad \ldots\ldots(3)$$

If $\alpha_1 > 1/\sqrt{2}$, then α_1 must be greater than either α_2 or α_3, and

$$f(\alpha_1) = 1 \quad (\alpha_1 > 1/\sqrt{2}). \qquad \ldots\ldots(4)$$

If $\alpha_1 < 1/\sqrt{3}$, at least one of the cosines α_2 or α_3 must be greater than α_1, and

$$f(\alpha_1) = 0 \quad (\alpha_1 < 1/\sqrt{3}). \qquad \ldots\ldots(5)$$

If $1/\sqrt{2} > \alpha_1 > 1/\sqrt{3}$, $f(\alpha_1)$ lies between 0 and 1 and is calculated as follows: from (1)

$$\alpha_2^2 + \alpha_3^2 = 1 - \alpha_1^2. \qquad \ldots\ldots(6)$$

Plot a point P whose Cartesian coordinates are α_2, α_3. From (6) P lies on the circumference of a circle of radius $(1-\alpha_1^2)^{\frac{1}{2}}$ as shown in Figure 3. By symmetry

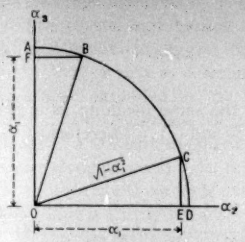


Figure 3.

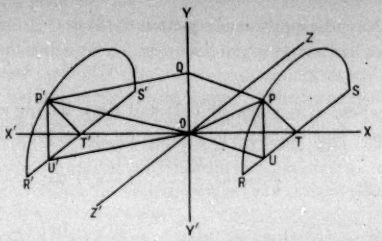


Figure 4.

only one quadrant need be considered. If P lies between B and C, $\alpha_1>\alpha_2$ and $\alpha_1>\alpha_3$, and

$$f(\alpha_1)=\text{arc BC/arc ABCD}=[\pi/2-2\cos^{-1}\{\alpha_1/(1-\alpha_1^2)^{\frac{1}{2}}\}]/(\pi/2)$$
$$=1-(4/\pi)\cos^{-1}\{\alpha_1(1-\alpha_1^2)^{-\frac{1}{2}}\}.\quad(1/\sqrt{2}>\alpha_1>1/\sqrt{3}).$$
$$\ldots\ldots(7)$$

From (1) and (2), remembering that $N\Delta V=1$,

$$J_r=-3J_s\int_1^0\alpha_1 f(\alpha_1)\,d\alpha_1,\qquad\ldots\ldots(8)$$

and from (4), (5), (7) and (8)

$$J_r=-3J_s\int_1^{1/\sqrt{2}}\alpha_1\,d\alpha_1-3J_s\int_{1/\sqrt{2}}^{1/\sqrt{3}}\alpha_1[1-(4/\pi)\cos^{-1}\{\alpha_1(1-\alpha_1^2)^{-\frac{1}{2}}\}]\,d\alpha_1.\ldots\ldots(9)$$

The first term of (9) gives 0·75, and the second term was integrated graphically and gave 0·087, so that

$$J_r=0{\cdot}837J_s.\qquad\ldots\ldots(10)$$

Gans (1932) using a more precise method not involving graphical integration, obtained $J_r=0{\cdot}831J_s$. The authors have not been able to apply Gans' method to the second case in which the alloy is magnetized perpendicular to OX, but the method given here can be adapted for their calculation.

Case 2. When a field is applied in a direction OY perpendicular to OX assume that at the remanence point the magnetic vector in all domains returns to the cube edge nearest to X'OX but in that sense which gives a positive component along OY. For a given value of α_1 we require the mean value of α_2. In the three dimensional Figure 4, OP and OP' represent the magnetic vector for a given value of α_1. P and P' lie on the arcs of two semicircles whose planes are parallel to the YZ plane. From this figure $\alpha_1=\cos\text{POX}=\cos\text{P'OX'}$; $\alpha_2=\cos\text{POY}=\cos\text{P'OY'}$.

Let PQ, P'Q be drawn perpendicular to OY, and PU, P'U' perpendicular to RS, R'S' respectively. Now $\text{OQ}=\text{OP}\alpha_2=\text{OP}'\alpha_2$; also

$$\text{OQ}=\text{OP}\sin\text{POX}\sin\text{PTU}=\text{OP}'\sin\text{P'OX'}\sin\text{P'T'U'}$$
$$=\text{OP}(1-\alpha_1^2)^{\frac{1}{2}}\sin\text{PTU}=\text{OP}'(1-\alpha_1^2)^{\frac{1}{2}}\sin\text{P''T'U'}.\qquad\ldots\ldots(11)$$

Since all values of angles PTU or P'T'U' are equally probable the required mean value of α_2 is

$$(2/\pi)(1-\alpha_1^2)^{\frac{1}{2}}. \qquad \ldots\ldots(12)$$

Now the equivalent equation to (8) is

$$J_r = -3J_s\int_1^0 \alpha_2 f(\alpha_1)\,d\alpha_1, \qquad \ldots\ldots(13)$$

where α_2 is given by (12) and $f(\alpha_1)$, having exactly the same meaning as in the previous calculation, is given by equations (2), (4), (5) and (7). Therefore

$$J_r = -J_s\frac{6}{\pi}\int_1^{1/\sqrt{2}}(1-\alpha_1^2)^{\frac{1}{2}}\,d\alpha_1$$
$$-J_s\frac{6}{\pi}\int_{1/\sqrt{2}}^{1/\sqrt{3}}(1-\alpha_1^2)^{\frac{1}{2}}\left(1-\frac{4}{\pi}\cos^{-1}\frac{\alpha_1}{(1-\alpha_1^2)^{\frac{1}{2}}}\right)d\alpha_1. \qquad \ldots\ldots(14)$$

The first term of the right-hand side of (14) is equal to $+0{\cdot}272$, and by graphical integration the second term is equal to $+0{\cdot}058$, so that $J_r = 0{\cdot}33J_s$.

REFERENCE

GANS, R., 1932, *Ann. Phys., Lpz.*, **15**, 28.

Magnetic Behaviour of Zinc and Cadmium at Low Temperatures

BY L. MACKINNON

The Royal Society Mond Laboratory, Cambridge

Communicated by D. Shoenberg; *MS. received* 26*th August* 1948

ABSTRACT. Following the recent discovery of the de Haas–van Alphen effect in zinc by Marcus at liquid hydrogen temperatures, the effect has been investigated at liquid helium temperatures using (for the most part) a torsion method similar to that used by Shoenberg in investigating bismuth. The effect in zinc takes the form of a variation of the susceptibility along the hexagonal axis which is periodic in the applied magnetic field. An attempt has been made to correlate the results with Landau's explicit formulae and to derive the various relevant parameters (such as effective masses, and degeneracy temperature) for the electrons causing the effect. Only qualitative agreement has been found, so that too much reliance should not be placed on the physical significance of the figures derived; however, it would appear that, as in the case of bismuth, the effect is caused by only a few times 10^{-6} electrons per atom. In spite of its very similar structure, cadmium did not show the effect; it had however, at liquid helium temperatures, a magnetic anisotropy some six times that at room temperature.

§1. INTRODUCTION

MARCUS (1947) has reported that, at 20° K., single crystals of zinc show a field dependence of magnetic susceptibility along the hexagonal axis. This effect is similar to that discovered in bismuth by de Haas and van Alphen (1930).

A theoretical explanation of this effect has been given for the case of an isotropic metal by Peierls (1933); this theory was extended to the particular case of bismuth by Blackman (1938) and Landau (see Shoenberg 1939). A detailed study of the effect in bismuth was made by Shoenberg (1939) and comparison made between

theory and experiment; although certain points of disagreement were found, it was possible to estimate the number, degeneracy temperature and effective masses of the electrons involved.

The crystal symmetries of bismuth (trigonal) and of zinc (hexagonal) are sufficiently alike to suggest that a similar theoretical approach might be applied to zinc. With a view to determining how far this is true, the measurements on zinc have been extended to the liquid helium temperature range. The principal method of investigation was the determination of the couple on the crystal when placed in a uniform magnetic field. This gives the difference between susceptibilities along perpendicular directions in the crystal, and this can be compared with the appropriate theoretical expression. Some preliminary experiments were also done using a Curie balance.

Since this investigation was started, a brief report of a similar study has been given by Sydoriak (1948), whose results have been briefly analysed by Onsager and Robinson (1948).

As cadmium has a crystal structure similar to that of zinc and as some similarity in properties might be expected, its magnetic anisotropy has also been studied by the torsion method down to 1·5° K.

§ 2. EXPERIMENTAL TECHNIQUE

The apparatus for the torsional investigation was essentially similar to that described by Shoenberg (1939). The torsion constant of the suspension (a phosphor-bronze wire) was determined by the timing of torsional oscillations; it was such that a couple of 1·81 dyne cm. gave a displacement of 1 cm. to the image observed in the travelling microscope; this image was 95 cm. from a mirror mounted on the suspension. Magnetic fields up to about 8,000 gauss were provided by an electromagnet, which was mounted on a graduated turntable so that its direction relative to the crystal could easily be varied.

The crystals used (two zinc and one cadmium) were grown by Bridgman's method (Bridgman 1925) modified to give a roughly spherical crystal.

The first zinc crystal (weighing 0·574 gm. after etching) was made from a spectroscopically pure sample (Hilger HS8392). The direction of the hexagonal (or z) axis was determined from the etch pits with a two-circle reflecting goniometer, and, on mounting, this axis was located perpendicular to the suspension, no account being taken of the orientation of the other axes. Torsional measurements with this crystal were largely in the nature of a "try-out" of the apparatus, but were accurate enough for subsequent analysis (§ 5).

The second zinc crystal (1·192 gm.) was made from a similar sample (Johnson Matthey JM1790). Practice with the etching technique enabled visual location of both the z axis and one set of diad axes; cleavage confirmed the z axis location and an x-ray photograph identified the diad axis. In the case of this crystal, the mounting on the suspension was better defined; two sets of measurements were carried out, both with the z axis perpendicular to the suspension, but the first set with the x axis also perpendicular and the second with it parallel to the suspension. The settings of the axes were probably within 3° of the intended orientations.

The cadmium crystal (1·850 gm.) was also made from a spectroscopically pure sample (Hilger HS8961). The direction of the hexagonal axis was identified visually after etching. This axis was then mounted at right angles to the suspension without identifying the other axes.

§3. EXPERIMENTAL RESULTS

(i) *Preliminary Measurements with a Curie Balance*

The first zinc crystal only was used for some preliminary measurements with a Curie balance. These measurements showed that the amplitude of the de Haas–van Alphen effect was, at liquid helium temperatures, at least 15 times as great with the field along the hexagonal axis as with it perpendicular; it should be noted that in bismuth it is parallel to the trigonal axis that the effect is absent. While this provided confirmation of Marcus' findings, and, as will be seen later (§5 (i)) provides some auxiliary information, the accuracy was not enough for detailed analysis of the finer features of the effect, and it was therefore decided to continue the investigation with the more convenient torsional arrangement.

(ii) *The Torsional Investigation*

With the crystal mounted as described in §2, if c be the couple on it in field H and ψ be the angle between field and hexagonal axis, it is easily shown that

$$\frac{c}{H^2 \cos\psi \sin\psi} = \frac{I_{\parallel}}{H_{\parallel}} - \frac{I_{\perp}}{H_{\perp}} \qquad \ldots\ldots(1)$$

where the $_{\parallel}$ and $_{\perp}$ suffixes denote the components of the intensity of magnetization I and the magnetic field H parallel and perpendicular to the z axis.

The most convenient way of presenting the results is thus in the form of graphs of $c/H^2 \cos\psi \sin\psi$ for various values of ψ and for various temperatures. Typical graphs, taken with the two settings of the second zinc crystal at 4·24° K., are shown in Figure 1. The general appearance of the graphs is very similar to those obtained by Shoenberg for bismuth, and, as in that case, the effect of changing the temperature is only to alter the amplitude of the oscillations without affecting the values of H at which their mid-points occur. Measurements at temperatures other than 4·24° K. were, for this crystal, only carried out for $\psi = 45°$ and the envelope curves of the oscillations at some of these other temperatures are shown on one of the graphs of Figure 1.

The first zinc crystal showed good agreement with the second for $\psi = 15°$, but diverged considerably at larger ψ, the amplitudes being somewhat greater than that for either setting of the second crystal. The exact crystal orientation was unknown, so that data for a full discussion of the discrepancy are lacking.

The interpretation of these zinc results is discussed in §5.

In the case of the cadmium crystal, the anisotropy was studied at 4·26° K. for some fifteen to twenty values of the field between 3,300 and 6,200 gauss for $\psi = 10°$ and $\psi = 80°$ without finding any trace of field dependence. The experiment was repeated at 1·50° K. and further measurements taken with $\psi = 5°$ and $\psi = 85°$ at fields up to 8,200 gauss; the findings remained negative although a change of 5% of the anisotropy with field would certainly have been observed. The magnitude of the anisotropy was identical at the two temperatures and some six times that at room temperature; approximate values obtained for $\chi_{\parallel} - \chi_{\perp}$ are $-0{\cdot}10 \times 10^{-6}$ per gm. at room temperature and $-0{\cdot}60 \times 10^{-6}$ per gm. at liquid helium temperatures.

§4. STATEMENT OF THEORY

The theoretical approach used by Blackman (1938) for bismuth is based on a representation of the surfaces of constant energy in k-space for the electrons concerned by ellipsoids located according to the crystal symmetry. Blackman

used three such ellipsoids for the bismuth case; he also adopted the simplest possible configuration of these ellipsoids (i.e. with axes along symmetry axes),

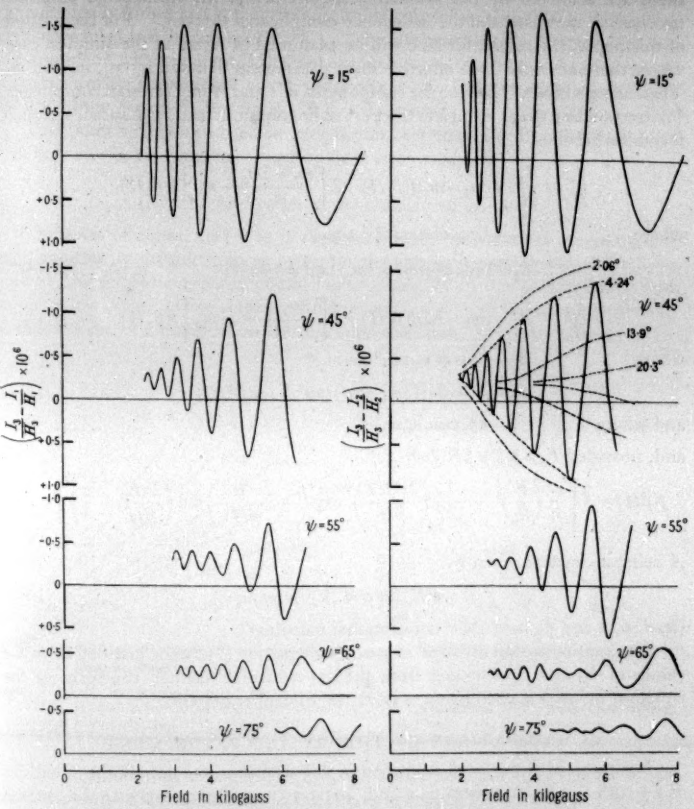


Figure 1. de Haas-van Alphen effect in zinc at 4·24° K.
(*a*) With *x*-axis ⊥ suspension. (*b*) With *x*-axis ∥ suspension.
Envelopes of curves at different temperatures shown by broken lines.
ψ is angle between field and hexagonal axis.

The experimental points have been omitted to avoid confusing the diagram, (e.g., there are 72 in the top-right hand curve), and for the most part the curves have been drawn through these points without smoothing so that deviations from the curves are very small. The accuracy with which points could be measured is much higher at high fields than at low; the light spot deflections, for example, were obtainable to 1 part in 1,000 at high fields and 1 part in 20 for low fields for $\psi=45°$.

which was later shown by the experimental work of Shoenberg (1939) to be an oversimplification. Landau put Blackman's work into an explicit form and also incorporated an extra parameter, which affected the configuration of the ellipsoids and was needed in order to interpret Shoenberg's results. The crystal symmetry

of zinc is sufficiently like that of bismuth to allow a representation similar to the simplest bismuth representation. Extra parameters could be included but, when these are acted on by the symmetry requirements, the number of ellipsoids necessary is increased and the issue much complicated thereby. For the purpose of this paper, the results for zinc will be examined in terms of the simplest case, where there are only three effective mass parameters denoted by m_1, m_2 and m_3. Then using suffixes 1, 2, 3 for the components of I and H in Cartesian coordinates, 1 corresponding to the x and 3 to the z crystallographic direction, Landau's explicit formulae become

$$\frac{I_3}{H_3}-\frac{I_1}{H_1}=(m_1-m_3)f(\beta_1 H)+2\left[\frac{m_1+3m_2}{4}-m_3\right]f(\beta_2 H), \quad \ldots\ldots(2)$$

where

$$\beta_1=\lambda\cos\psi(m_1\tan^2\psi+m_3)^{\frac{1}{2}}, \quad \ldots\ldots(3)$$

$$\beta_2=\lambda\cos\psi\{\tfrac{1}{4}(m_1+3m_2)\tan^2\psi+m_3\}^{\frac{1}{2}}; \quad \ldots\ldots(4)$$

$$\frac{I_3}{H_3}-\frac{I_2}{H_2}=(m_2-m_3)f(\beta_1' H)+2\left[\frac{3m_1+m_2}{4}-m_3\right]f(\beta_2' H), \quad \ldots\ldots(5)$$

where

$$\beta_1'=\lambda\cos\psi(m_2\tan^2\psi+m_3)^{\frac{1}{2}}, \quad \ldots\ldots(6)$$

$$\beta_2'=\lambda\cos\psi\{\tfrac{1}{4}(3m_1+m_2)\tan^2\psi+m_3\}^{\frac{1}{2}}, \quad \ldots\ldots(7)$$

and where

$$\lambda=\boldsymbol{e}\hbar/\boldsymbol{c}(m_1m_2m_3)^{\frac{1}{2}}, \quad \ldots\ldots(8)$$

and, provided $E_0\gg\boldsymbol{k}T\gg\beta H/2\pi^2$,

$$f(\beta H)=A\left[\frac{\pi^2}{6}\left(\frac{\boldsymbol{k}}{E_0}\right)^{\frac{1}{2}}-\frac{1}{\sqrt{T}}\left(\frac{2\pi^2\boldsymbol{k}T}{\beta H}\right)^{\frac{1}{2}}\exp\left(-\frac{2\pi^2\boldsymbol{k}T}{\beta H}\right)\sin\left(\frac{2\pi E_0}{\beta H}-\frac{\pi}{4}\right)\right], \quad \ldots\ldots(9)$$

A being a constant given by

$$A=\boldsymbol{e}^2E_0/[\pi^4\boldsymbol{c}^2\hbar(2\boldsymbol{k})^{\frac{1}{2}}(m_1m_2m_3)^{\frac{1}{2}}], \quad \ldots\ldots(10)$$

($\boldsymbol{k}$, $\hbar$, $\boldsymbol{c}$, $\boldsymbol{e}$ and E_0 have their conventional meanings).

The approximation involved in assuming equation (9) can be justified from the values of E_0 and β estimated from the experimental results; the form of the equation before approximation is given by Shoenberg (1939).

§ 5. COMPARISON OF THEORY AND EXPERIMENT

The purpose of the work described in this section is to compare the equations of § 4 with the experimental data and, as far as possible, to evaluate the various parameters which enter into these equations and describe the electrons giving rise to the de Haas–van Alphen effect. The value of n, the number of electrons per atom responsible for the effect is also derived.

(i) *Information given by the Curie Balance Experiment mentioned in* § 2

The absence, within experimental error, of a field dependence of susceptibility perpendicular to the hexagonal axis gives a little preliminary information on the relative sizes of the m's of § 4. Assuming for simplicity that $m_1=m_2=m_\perp$ and $m_3=m_\parallel$ and writing $x_\perp=m_\perp/m_0$ and $x_\parallel=m_\parallel/m_0$ where m_0 is the true electron mass, it can be shown that

$$\frac{\text{Amp}_\perp}{\text{Amp}_\parallel}=\left(\frac{x_\perp}{x_\parallel}\right)^{\frac{1}{2}}\exp\left\{-\frac{2\pi^2\boldsymbol{k}Tcm_0}{H\boldsymbol{e}\hbar}x_\parallel\left[\left(\frac{x_\perp}{x_\parallel}\right)^{\frac{1}{2}}-\left(\frac{x_\perp}{x_\parallel}\right)\right]\right\} \quad \ldots\ldots(11)$$

where "Amp" denotes the amplitude of the effect, the magnetic field being in the direction implied by the subscript. Since the ratio of the amplitudes is less than 1, it is immediately clear from (11) that $x_\perp/x_\parallel$ must also be less than 1. Now the experimentally found fact is that the amplitude ratio is less than 0·07 at the highest H (9,000 gauss) and the lowest T (2° K.). The information obtained from this fact in conjunction with equation (11) is not as much as might at first be hoped because there is a drastic dependence on the value of $x_\parallel$. If $x_\parallel = 1$, the experimental facts are fitted by $x_\perp/x_\parallel < 0{\cdot}8$; but for $x_\parallel = 0{\cdot}1$, it is necessary that $x_\perp/x_\parallel < 10^{-4}$.

This point will be returned to later in this section (see (viii) below) after the results from the torsional investigation have been analysed.

(ii) *The Mid-points of the Oscillations at low* ψ

The form of equations (2) to (9) is considerably simplified by assuming that m_3 is sufficiently greater than m_1 or m_2 for the approximation $\beta_1 = \beta_2 = \beta_1' = \beta_2'$ to hold at low ψ. The experimentally found similarity in the susceptibility-field curves at low ψ for the two settings of the second crystal and the arbitrary setting of the first provide justification for this approximation.

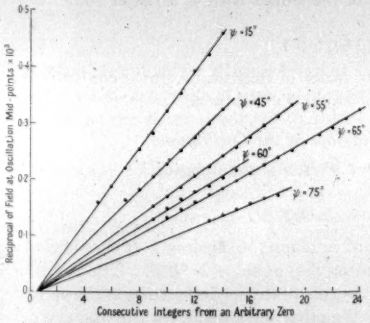


Figure 2. $1/H$ for mid-points of oscillations plotted against consecutive integers; these integers have been chosen to make the curves for each ψ intersect as shown.

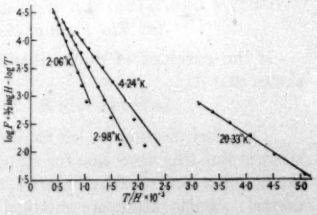


Figure 3. Graph of $\log F + 3/2 \log H - \log T$ plotted against T/H; logarithms to base 10 have been used for convenience.

Now, examination of equations (2) to (9) shows that, assuming the β's equal, the mid-points of the oscillations in the curves of Figure 1 should occur whenever

$$2\pi E_0/\beta H = r\pi + \pi/4 \qquad \ldots\ldots(12)$$

where r is an integer, irrespective of the temperature. Thus plotting the values of $1/H$ for the mid-points against a series of consecutive integers should give a straight line of slope $\beta/2E_0$. This has been done and Figure 2 shows the results for the first setting of the second crystal for various values of ψ; similar results were obtained from the other measurements, but are not illustrated. As mentioned in §3, the experimentally found mid-points are independent of temperature, as predicted by theory.

It may be noted that $\pi/2$ fits the experimental results better than $\pi/4$ in equation (12); Shoenberg (1939) also found a discrepancy here in the case of bismuth, preferring $-\pi$ to $\pi/4$. There is no obvious explanation of this discrepancy, but

the possibility that it arises from experimental errors is not excluded in either case.

Now, for low ψ, equation (3) etc., gives approximately

$$\beta = \lambda \cos\psi m_3^{\frac{1}{2}} \qquad \ldots\ldots(13)$$

irrespective of the location of the x crystal axis.

Thus the slope of the graph of Figure 2 for low ψ gives $\lambda m_3^{\frac{1}{2}}/2E_0$; this reduces algebraically to $\beta_0/\boldsymbol{k}(x_1x_2)^{\frac{1}{2}}T_0$ where β_0 is the Bohr magneton, T_0 the degeneracy remperature ($E_0/\boldsymbol{k}$), and x_1, x_2 denote m_1/m_0 and m_2/m_0 (m_0 being the electron mass) tespectively. Appropriate extrapolation of the slopes to zero ψ gives the quantity

$$P = (x_1x_2)^{\frac{1}{2}}T_0. \qquad \ldots\ldots(14)$$

The measurements at 4·24° K. gave $P = 1{\cdot}96 \pm 0{\cdot}02$ centigrade degrees; since the mid-points do not depend on temperature (§3), measurements at other temperatures give the same result.

(iii) *The Variation with ψ of the Slopes of the Lines in Figure 2*

This was found difficult to reconcile with the simple theory, suggesting, among other things, that at high ψ, the β's could no longer be treated as equal (see also (v) below). Although, from the variation of the slopes with ψ, a rather rough estimation gives

$$(m_1 + m_2)/m_3 \simeq 0{\cdot}1, \qquad \ldots\ldots(15)$$

this result is certainly not more than an order of magnitude, but it confirms the reasonableness of the approximation of equal β's at low ψ.

(iv) *The Form of the Envelope of the Oscillations*

If the envelope of the oscillations is $F(H, T)$, examination of equation (9) shows that

$$\ln F + (3/2)\ln H - \ln T = -2\pi^2\boldsymbol{k}T/\beta H + \text{constant}. \qquad \ldots\ldots(16)$$

Therefore plotting the left hand side of equation (16) against T/H should give a straight line (the same line for all temperatures) of slope $2\pi^2\boldsymbol{k}/\beta$. The result of doing this is illustrated by Figure 3, for the case of the second setting of the second crystal; similar results are obtained for the other setting and for the first crystal. It is immediately apparent that there is something wrong here; the quantity $2\pi^2\boldsymbol{k}/\beta$ appears as a function of temperature, which would appear to be unlikely in view of the constancy of β/E_0 noted at (ii) above. Thus, while the variation of F with H appears nearly correctly predicted, that of F with T does not. In fact the experimental findings show that the left hand side of (15) cannot be expressed as a function of T/H. A similar discrepancy, although not quite as marked, was noted by Shoenberg for bismuth. This discrepancy does not seem to be a consequence of any of the approximations.

However, from the slopes of Figure 3 and similar graphs, the quantity β can be determined for each temperature and setting. Using equations (12) and (8), one can then calculate

$$Q = (x_1x_2)^{\frac{1}{2}} \qquad \ldots\ldots(17)$$

and then, using equation (14), estimate T_0. In view of the discrepancy just noted, the physical interpretation of these parameters must be regarded with some reserve. The numerical values of Q and T_0 obtained are listed in Table 1. Since, as already mentioned, the assumption of equal β's is valid only for fairly low ψ, graphs for $\psi > 45°$ were not considered in this analysis.

(v) "*Beats*" *in the Oscillations*

Further evidence that the β's cannot be treated as equal for large ψ is provided by the occurrence of "beats" in certain cases. An illustration of this is shown in the graph of Figure 1 for $\psi = 65°$ and the second setting; these "beats" are of the type associated with the superposition of periodic terms of slightly differing periodicities. Assuming slightly unequal β's, it is then possible to estimate roughly from the period of the "beats" that

$$|(m_1 - m_2)/m_3| \simeq 0{\cdot}1. \qquad \ldots\ldots(18)$$

Since the estimates (15) and (18) are only orders of magnitude, nothing definite can be deduced from them about the relative sizes of m_1 and m_2 but the estimates are not inconsistent with the interpretation that either m_1 or m_2 is considerably the larger. This point will be discussed again later.

(vi) *The Absolute Value of the Amplitude of the Oscillations*

From the numerical value of this quantity, using the equations of §4, and assuming the β value determined as in (iv) above, an estimate is obtained of the quantity

$$R = (x_3/x_1x_2)^{\frac{1}{2}} T_0. \qquad \ldots\ldots(19)$$

This has been done for all cases listed in Table 1 by measuring the amplitude of the field dependence graph at some appropriate field strength. Because the envelope has almost the shape predicted by theory, the particular field chosen is relatively

Table 1. Summarized Results of Calculations

(1) Crystal	(2) Setting		(3) T	(4) $Q(=x_1x_2)^{\frac{1}{2}} \times 10^2$	(5) T_0	(6) $R=(x_3/x_1x_2)^{\frac{1}{2}}T_0 \times 10^{-3}$	(7) x_3	(8) $n \times 10^6$
First	30°	—	1·39	2·8	68	3·4	1·9	3·6
			2·08	2·0	98	4·1	0·7	2·7
			3·00	1·38	139	4·6	0·21	1·7
			4·25	1·08	178	4·6	0·08	1·2
Second	45°	First	2·12	3·1	64	2·4	1·3	3·0
			2·98	2·5	80	3·5	1·16	3·2
			4·24	1·8	111	3·8	0·4	2·1
		Second	2·06	2·5	78	2·5	0·7	2·3
			2·98	1·9	104	3·1	0·3	1·9
			4·24	1·5	134	3·8	0·17	1·6
			20·33	0·7	265	4·7	0·018	0·7
	15°	First	4·24	1·05	187	3·1	0·03	0·8
		Second	4·24	1·05	187	3·6	0·04	0·9

(3) absolute temperature T in °K. (4) effective mass $Q=(x_1x_2) =(m_1m_2)^{\frac{1}{2}}/m_0$ in electron masses $\times 10^2$ (see equation (17) of text). (5) apparent degeneracy temperature T_0 in °K. (6) the quantity $R=(x_3/x_1x_2)^{\frac{1}{2}}T_0 \times 10^{-3}$ (see equation (19) of text); (7) effective mass $x_3=m_3/m_0$ in electron masses; (8) the apparent number of electrons per atom, n, causing the effect $\times 10^6$.

unimportant. That those R values vary with temperature is due to the inadequacy in the theory pointed out in (iv) above. That R varies with setting also does not agree with the simple theory, but this discrepancy may have its origin elsewhere than in the temperature-dependence discrepancy.

(vii) *The Parameter* $x_3 = m_3/m_0$

From equations (14), (17) and (19), it follows that

$$x_3 = P^2Q^4R^2. \qquad \ldots\ldots(20)$$

The values of x_3 so obtained are incorporated in Table 1.

(viii) *The Relative Sizes of* x_1, x_2 *and* x_3

Since $(x_1+x_2)/2x_3$ must always be greater than $(x_1x_2)^{\frac{1}{2}}/x_3$, it is evident that it becomes difficult to reconcile the estimates of $(x_1x_2)^{\frac{1}{2}}/x_3$ obtainable from Table 1 at 4·24° K. and 20·33° K. with the estimate of $(x_1+x_2)/2x_3$ obtainable from equation (15). At the lower temperatures, however, the estimates are consistent with either x_1 or x_2 being much larger (~10 times) than the other; this tallies with equation (18), derived from the "beats".

The experimental evidence, mentioned in § 3 (i) and discussed in § 5 (i), about the absence of a field dependence perpendicular to the hexagonal axis, also leads to difficulties at the higher temperatures. While the figures at 2·08° K. are quite reconcilable with equation (11), those for the first crystal at 4·25° K. are not; the same discrepancy occurs at this temperature for the second setting of the second crystal at $\psi = 45°$ but not for the first setting.

Thus there is some suggestion that the estimates of the parameters obtained from the measurements at lower temperatures are more self-consistent than those at the higher temperature.

(ix) *The Number of Electrons per Atom, n, contributing to the Effect*

This is given by Shoenberg (1939)

$$n = (2E_0)^{\frac{3}{2}}(m_1m_2m_3)^{\frac{1}{2}}V/\pi^2h^3N \qquad \ldots\ldots(21)$$

where V is the atomic volume and N is Avogadro's number.

The estimates of this quantity have been added to Table 1. Although these figures, like the others, must be regarded with reserve owing to the discrepancy found between theory and experiment, it is probably reasonable to assume that the number of electrons involved is of the order of a few times 10^{-6} per atom; this agrees with Shoenberg's estimate in the case of bismuth.

§ 6. DISCUSSION OF RESULTS

It is quite clear that a simple interpretation of the experimental findings in terms of the theory of § 4 is not possible. It is true that the value of x_3, for example, are open to suspicion since high powers of the observed quantities are used in their derivation (equation (20)); however the systematic nature of the variation with temperature of x_3 and the other parameters listed in Table 1 (all of which should be constants) suggests that this variation is not the result of random experimental errors but is due rather to the inadequacy of the theory. Clearly, until a more adequate theory taking account of these variations is developed, it is difficult to know how much physical significance can be attached to the estimates in Table 1. Probably they do give some idea of the orders of magnitude and, if the considerations of § 5 (viii) above are to be believed, the estimates based on the measurements at the lower temperature may give the more reliable picture. To obtain a better picture, it may be necessary (*a*) to examine the fundamental assumptions of the theory (which include an assumption that the number of electrons involved

does not vary with temperature) and check their validity in order to remove discrepancies such as that noted in § 5 (iv), and also (*b*) to examine whether the assumption of only three ellipsoids in *k*-space is not a serious oversimplification (which is probably the case) and, if it appears to be so, to obtain further experimental measurements to determine a distribution of ellipsoids which will tally with the facts. However it should be noted that no redistribution of the ellipsoids can remove the temperature discrepancy noted in (*a*).

One possible clue in seeking a reformulation of the theory is provided by Marcus' additional observation that, at temperatures above those at which the de Haas–van Alphen effect appears, $\chi_{\parallel}-\chi_{\perp}$ in zinc varies with temperature in an unusual manner.* This behaviour is not explained by theory and, although it may have nothing to do with the de Haas–van Alphen effect, it is an anomaly which may possibly require explanation before a more complete theory of the de Haas–van Alphen effect in zinc can be formulated.

The results of Table 1 do not entirely agree with the figures given so far by Onsager and Robinson—the only notable discrepancy is that they estimate $x_3 \sim 20$—but, until further details of their work are available, it is hardly possible to discuss this point.

In the case of cadmium, it has been shown that, at 1·50° K. and 8,000 gauss, the amplitude of the oscillation, if it occurs, is certainly less than $0{\cdot}03\times10^{-6}$ units of susceptibility per gramme. Using $m_{\parallel}$ and $m_{\perp}$ as in § 5 (i), from the equations of § 4, this amplitude is given by theory as

$$3(m_{\perp}-m_{\parallel})\frac{A}{\sqrt{T}}\left(\frac{2\pi^2\boldsymbol{k}T}{\beta H}\right)^{\frac{3}{2}}\exp\left(-\frac{2\pi^2\boldsymbol{k}T}{\beta H}\right)\text{units/cm}^3. \quad \ldots\ldots(22)$$

Substituting numerical values, it follows that, if there are any electrons in cadmium corresponding to those giving rise to the de Haas–van Alphen effect in zinc, they can be described by the inequality

$$\{B^{\frac{3}{2}}E_0 m_{\perp}^{\frac{1}{2}}(m_{\perp}-m_{\parallel})/m^{\frac{1}{2}}\}\exp(-Bm_{\perp})<0{\cdot}58 \quad \ldots\ldots(23)$$

where B is a constant of numerical value $1{\cdot}13\times10^{29}$; this inequality is based on the absence of field dependence at $\psi=5°$ and assumes $m_{\parallel}\gg0{\cdot}008m_{\perp}$; a similar inequality can be deduced from the absence of field dependence for large ψ.

§ 7. CONCLUSIONS

The conclusions to be drawn from these experiments and their analysis are as follows :

(1) The de Haas–van Alphen effect in zinc, as in bismuth, is caused by a small number of electrons, probably of the order of a few times 10^{-6} per atom.

(2) Existing theory is not adequate for more than a qualitative picture of the effect, breaking down particularly in the matter of its temperature dependence.

(3) Although cadmium is very similar to zinc, it does not show the effect.

ACKNOWLEDGMENTS

The author desires to thank Dr. D. Shoenberg for his kind assistance and advice in this work and for drawing his attention to the possibility of the problem. He

* During this investigation, it was possible to confirm qualitatively Marcus' observation; this was done by allowing the crystal to warm up from liquid air temperatures when the liquid had evaporated and following the light-spot deflection at constant field.

also wishes to thank Dr. C. L. Smith for taking and analysing the X-ray photograph referred to in § 2, for advice in crystal growing and for the assistance given by him and his co-workers in the earlier goniometry. The work was carried out while in receipt of a D.S.I.R. maintenance grant.

REFERENCES

BLACKMAN, M., 1938, *Proc. Roy. Soc.* A, **166,** 1.
BRIDGMAN, P. W., 1925, *Proc. Amer. Acad. Arts and Sci.*, **60,** 303.
DE HAAS, W. J., and VAN ALPHEN, P. M., 1930, *Leiden Comm.*, 212 *a*.
MARCUS, J. A., 1947, *Phys. Rev.*, **71,** 559.
ONSAGER, L., and ROBINSON, J. E., 1948, *Bull. Amer. Phys. Soc.*, **23,** No. 3, p. 40.
PEIERLS, R., 1933, *Z. Phys.*, **81,** 186.
SHOENBERG, D., 1939, *Proc. Roy. Soc.* A, **170,** 341.
SYDORIAK, S. G., 1948, *Phys. Rev.*, **73,** 1247.

On the Dielectric Properties of a Gas Discharge

BY E. E. SALPETER* AND R. E. B. MAKINSON
School of Physics, University of Sydney

Communicated by L. G. H. Huxley; MS. received 2nd July 1948

ABSTRACT. Formal expressions are derived for the space currents and electrode currents produced by a sinusoidal potential difference of angular frequency ω applied between a pair of electrodes at a distance d apart and immersed in a gas discharge. It is shown that the ionized gas behaves like an isotropic medium of dielectric constant $\{1-4\pi ne^2/m\omega^2\}$ and zero conductivity only in the limiting case of $\beta \equiv v/\omega d \ll 1$ where v is the R.M.S. velocity of the free electrons in the gas. The variation of the equivalent dielectric constant and conductivity of the gas with β are sketched. Extensions of Green's reciprocation theorem and of a theorem on induced currents given by W. Shockley are derived in the appendix.

§ 1. INTRODUCTION

MEASUREMENTS have been taken in this laboratory of the impedance presented to a high-frequency circuit at about 1,000 Mc/s. by a pair of electrodes projecting into a gas discharge, their separation being much less than one wavelength, with the object of deriving information about the electron concentration and temperature in the discharge. It became necessary in this work to estimate the effect of the random velocity of the electrons on the electrode currents. Such random velocities may be quite high, of the order of 10^8 cm/sec. corresponding to a temperature of about 30,000° K., say. The transit time, τ, taken by an electron to cross the region near the electrodes in which the electric field is appreciable is then quite small (less than 10^{-8} sec., say, if the spacing of the electrodes is about 1 cm.) and may be of the same order as the period of the field (10^{-9} sec. at the above frequency). In such cases, even if the effect of collisions can be neglected, it is shown below that the electrodes do not in general behave as if immersed in a medium with zero conductivity and with a dielectric constant given by the well known expression $1-4\pi ne^2/m\omega^2$ where n is the electron density as is the case when τ is much longer than the period. The discrepancy

* Now at Department of Mathematical Physics, University of Birmingham.

becomes marked only if the extent of the field is sufficiently small at a given frequency (thus being unimportant in the ionosphere) or if the frequency is sufficiently low for a given extent of the field and electron temperature. Qualitatively such a discrepancy is to be expected when the electrons have not time during their transit through the field to acquire from or restore to the field as much energy as in the case of slow transit. Thus the effect of the electrons on the equivalent dielectric constant of the space is reduced and some energy acquired by electrons in the field may be conveyed elsewhere.

Formal expressions are derived below for the space currents carried by the electrons and for the electrode currents taking the transit time τ into consideration.

§ 2. FORMAL EXPRESSIONS FOR THE SPACE CURRENTS

We shall consider a pair of electrodes of any shape at a distance of the order of magnitude of d from each other immersed in an ionized gas with arbitrary velocity distributions for the electrons and ions. Let $n(\mathbf{r}, \mathbf{v})$ be the number of free electrons contained in unit volume in the neighbourhood of the space-point $\mathbf{r}$ with velocities contained in unit volume of velocity space near the velocity $\mathbf{v}$. Let a sinusoidal potential difference represented by $Ae^{j\omega t}$ be applied between the electrodes and let the electric field $\mathbf{E}(\mathbf{r}, t)$ produced at a point $\mathbf{r}$ in the gas be $\Xi(\mathbf{r})Ae^{j\omega t}$. In general $\Xi(\mathbf{r})$ may be complex (i.e. field and potential difference may not be in phase) and different from the real quantity $\Xi_0(\mathbf{r})$ for the case of the same electrodes in vacuum. Ξ will be very much less than d^{-1} outside a region of extent of the order of d.

We shall restrict ourselves to cases where the following simplifications hold: (i) The extent of the field is much less than the mean free path of the electrons so that the effect of gas collisions can be neglected. (ii) d is much less than $2\pi c/\omega$ so that retardation effects can be neglected. (iii) The amplitude A is sufficiently small for the changes of position, velocity and density due to the applied field to be much less than d, v and n respectively. We also make the further approximations: (iv) The contribution of the positive ions to the space currents is neglected because of their larger mass. (v) The effect of static space charges is neglected. (vi) The effect on the electrode currents of direct collisions of electrons or positive ions with the electrodes is neglected.

At a particular instant t let $\Delta\mathbf{r}(\mathbf{r}, \mathbf{v}, t)$ and $\Delta\mathbf{v}(\mathbf{r}, \mathbf{v}, t)$ be the changes of position and velocity respectively, due to the field, for electrons which would have arrived at the space point $\mathbf{r}$ with velocity $\mathbf{v}$ in the absence of the field. The change in $n(\mathbf{r}, \mathbf{v})$ is then $\Delta n(\mathbf{r}, \mathbf{v}, t) = -\operatorname{div}\{n(\mathbf{r}, \mathbf{v})\Delta\mathbf{r}(\mathbf{r}, \mathbf{v}, t)\}$. Let $\Delta\mathbf{i}(\mathbf{r}, \mathbf{v}, t)dv$ [(3)] be the change in the space current density near the point $\mathbf{r}$ due to electrons with velocities in the small volume of velocity space dv [(3)]. If e and m be the charge and mass respectively of an electron, then

$$\Delta\mathbf{i}(\mathbf{r}, \mathbf{v}, t) = e\{n(\mathbf{r}, \mathbf{v})\Delta\mathbf{v}(\mathbf{r}, \mathbf{v}, t) - \mathbf{v}\,.\,\operatorname{div}[n(\mathbf{r}, \mathbf{v})\Delta\mathbf{r}(\mathbf{r}, \mathbf{v}, t)]\}. \quad \ldots\ldots(1)$$

Integrating the acceleration given to an electron by the field we have

$$\Delta\mathbf{v}(\mathbf{r}, \mathbf{v}, t) = \frac{Ae}{m}\int_{-\infty}^{t} \Xi(\mathbf{r} + [t' - t]\mathbf{v})e^{j\omega t'}dt' = \frac{Ae}{m}\mathbf{L}(\mathbf{r}, \mathbf{v})e^{j\omega t} \quad \ldots\ldots(2)$$

where

$$\mathbf{L}(\mathbf{r}, \mathbf{v}) = \int_{-\infty}^{0} \Xi(\mathbf{r} + t\mathbf{v})e^{j\omega t}\,dt. \quad \ldots\ldots(2a)$$

Similarly, integrating the change in velocity $\Delta\mathbf{v}$, we have

$$\Delta\mathbf{r}(\mathbf{r}, \mathbf{v}, t) = \int_{-\infty}^{t} \frac{Ae}{m} \mathbf{L}(\mathbf{r} + [t' - t]\mathbf{v}, \mathbf{v}) e^{j\omega t'} dt' = \frac{Ae}{m} \mathbf{\Lambda}(\mathbf{r}, \mathbf{v}) e^{j\omega t} \quad \ldots\ldots(3)$$

where

$$\mathbf{\Lambda}(\mathbf{r}, \mathbf{v}) = \int_{-\infty}^{0} \mathbf{L}(\mathbf{r} + t\mathbf{v}, \mathbf{v}) e^{j\omega t} dt. \quad \ldots\ldots(3\,a)$$

Hence the total space current density $\mathbf{I}(\mathbf{r}, t)$ near the point $\mathbf{r}$ is given by integrating (1) over the velocity distribution (which might for example be Maxwellian).

$$\mathbf{I}(\mathbf{r}, t) = \frac{Ae^2}{m} e^{j\omega t} \int \{n(\mathbf{r}, \mathbf{v})\mathbf{L}(\mathbf{r}, \mathbf{v}) - \mathbf{v} \cdot \mathrm{div}[n(\mathbf{r}, \mathbf{v})\mathbf{\Lambda}(\mathbf{r}, \mathbf{v})]\} dv^{(3)}. \quad \ldots\ldots(4)$$

In general $\mathbf{L}$ and $\mathbf{\Lambda}$ are complex and the space current $\mathbf{I}(\mathbf{r}, t)$ not parallel to nor in quadrature with the electric field $\mathbf{E}(\mathbf{r}, t)$. Thus the medium does not in general behave like a loss-free, isotropic dielectric. However, $\mathbf{I}(\mathbf{r}, t)$ is still sinusoidal and proportional to $E(\mathbf{r}, t)$ and we can therefore define a generalized "vector dielectric constant" $\mathbf{K}(\mathbf{r})$ and a electric displacement $\mathbf{D}(\mathbf{r}, t)$ as follows:

$$\mathbf{D}(\mathbf{r}, t) \equiv \mathbf{K}(\mathbf{r}) E(\mathbf{r}, t) \equiv \mathbf{E}(\mathbf{r}, t) + (4\pi/j\omega)\mathbf{I}(\mathbf{r}, t). \quad \ldots\ldots(5)$$

Even when $\Xi(\mathbf{r})$ is known throughout the medium the evaluation of the integrals in $(2\,a)$ and $(3\,a)$, and hence of the vector dielectric constant $\mathbf{K}(\mathbf{r})$, is in general very involved. Furthermore, $\Xi(\mathbf{r})$ depends on the space currents $\mathbf{I}(\mathbf{r}, t)$ throughout the medium as well as on the electrode spacing unless all $n(\mathbf{r}, \mathbf{v})$ are very small. For the special cases of very short and very long transit time, however, approximations for the integrals in $(2\,a)$ and $(3\,a)$ can be derived, making some general conclusions possible. This is treated in the next section.

§ 3. APPROXIMATIONS FOR LONG AND SHORT TRANSIT TIME

Since $\Xi \ll d^{-1}$ outside a region of extent of the order of d we may call $\tau = d/v$ (where v is the R.M.S. thermal velocity for the electrons in the gas) the transit time and we set

$$\beta \equiv v/\omega d. \quad \ldots\ldots(6)$$

We will treat separately the cases where $\beta \ll 1$ and where $\beta \gg 1$, i.e. where the transit time is much longer and much shorter respectively than the period of the field.

(a) *Long transit time.*

Here $\beta \ll 1$ and hence Ξ in $(2\,a)$ is a slowly varying function of t compared with the $e^{j\omega t}$ term. Successive integration by parts yields an asymptotic expansion for $(2\,a)$ in the form

$$\mathbf{L}(\mathbf{r}, \mathbf{v}) = \frac{1}{j\omega}\left\{\Xi(\mathbf{r}) + j\frac{\Xi'(\mathbf{r})v}{\omega} - \frac{\Xi''(\mathbf{r})v^2}{\omega^2} + \ldots.\right\}, \quad \ldots\ldots(7)$$

where $\Xi^{(s)}(\mathbf{r})$ is the sth space derivative of $\Xi(\mathbf{r})$ in the direction of $\mathbf{v}$ taken at the point $\mathbf{r}$. Similarly we have

$$\mathbf{\Lambda}(\mathbf{r}, \mathbf{v}) = -\frac{1}{\omega^2}\left\{\Xi(\mathbf{r}) + 2j\frac{\Xi'(\mathbf{r})v}{\omega} - \ldots.\right\}. \quad \ldots\ldots(8)$$

Substituting (7) and (8) in (1) we have (writing n, Ξ etc. for $n(\mathbf{r}, \mathbf{v})$, $\Xi(\mathbf{r})$ etc.)

$$\Delta\mathbf{i}(\mathbf{r}, \mathbf{v}, t) = \frac{Ae^2}{mj\omega} e^{j\omega t}\left\{n\Xi + j\left[\frac{n\Xi' v}{\omega} + \frac{\operatorname{div}(n\Xi)\mathbf{v}}{\omega}\right] - \left[\frac{n\Xi'' v^2}{\omega^2} + \frac{2\operatorname{div}(n\Xi')\mathbf{v}v}{\omega^2}\right] + \ldots\right\}. \quad \ldots\ldots(9)$$

If $\Xi(\mathbf{r})$ and $n(\mathbf{r}, \mathbf{v})$ do not vary too abruptly throughout space then Ξ', Ξ'', $\operatorname{div}(n\Xi)$, $\operatorname{div}(n\Xi')$ will be of the same order of magnitude as Ξd^{-1}, Ξd^{-2}, $n\Xi d^{-1}$, $n\Xi d^{-2}$ respectively. The successive terms in (9) thus form an expansion in ascending powers of β. For small β this expansion converges rapidly but for $\beta \gtrsim 1$ the expansion (9) is invalid. In calculations where the effect of the finite transit time is neglected only the first term of (9) occurs. The contribution of this term to the space current density and hence to $\mathbf{K}(\mathbf{r})$ as defined in (5) is in a direction parallel to the field $\mathbf{E}(\mathbf{r}, t)$ and independent of $\mathbf{v}$. Hence if the higher terms are omitted the electron gas behaves as an isotropic, loss-free dielectric of dielectric constant $K(\mathbf{r})$ given by the well known expression

$$K(\mathbf{r}) = 1 - \frac{4\pi e^2}{m\omega^2}\int n(\mathbf{r}, \mathbf{v})dv^{(3)}. \quad \ldots\ldots(10)$$

Since the phase of $\Xi(\mathbf{r})$ will in general not vary rapidly with position (see § 4) the second term in (9) represents a space current approximately in phase with the field $\mathbf{E}$ of order of magnitude β times that of the first (polarization current) term and the third term a space current approximately in quadrature with $\mathbf{E}$ of order of magnitude β^2 times that of the first term.

(b) *Short transit time.*

Here $\beta \gg 1$ and hence Ξ in (2 *a*) is appreciable over a small fraction of a cycle only and we can expand the $e^{j\omega t}$ term as a power series in ωt. This yields

$$\mathbf{L}(\mathbf{r}, \mathbf{v}) = \int_{-\infty}^{0} \Xi(\mathbf{r} + t\mathbf{v})dt + j\omega\int_{-\infty}^{0} \Xi(\mathbf{r} + t\mathbf{v})t\,dt + \ldots. \quad \ldots\ldots(11)$$

and a similar expansion for $\mathbf{\Lambda}(\mathbf{r}, \mathbf{v})$. When these expressions for $\mathbf{L}$ and $\mathbf{\Lambda}$ are substituted in (1) an asymptotic expansion is obtained for $\Delta\mathbf{i}(\mathbf{r}, \mathbf{v}, t)$. Since $\Xi(\mathbf{r})$ is appreciable over a region of extent d only the first few terms form an expansion in descending powers of β, the leading term $\Delta\mathbf{i}^{(1)}(\mathbf{r}, \mathbf{v}, t)$ being given by

$$\Delta\mathbf{i}^{(1)}(\mathbf{r}, \mathbf{v}, t) = \frac{Ae^2}{m} e^{j\omega t}\left\{n\int_{-\infty}^{0} \Xi(\mathbf{r} + t\mathbf{v})dt - \mathbf{v}.\operatorname{div}\left[n\int_{-\infty}^{0}\int_{-\infty}^{0} \Xi(\mathbf{r} + [t + t']\mathbf{v})dt\,dt'\right]\right\}. \quad \ldots\ldots(12)$$

The two terms in (12) are in general of opposite sign and each is of the order of magnitude of $j\beta^{-1}$ times the leading (polarization current) term in (9). Hence in this case ($\beta \gg 1$) the electron gas does not behave at all like a loss-free isotropic dielectric; the space currents depend very markedly, both in magnitude and direction, on $\Xi(\mathbf{r})$ and $n(\mathbf{r}, \mathbf{v})$ throughout the medium and are approximately in phase with the field $\mathbf{E}$ (since the phase of $\Xi(\mathbf{r})$ again does not vary rapidly). A generalized complex vector dielectric constant $\mathbf{K}(\mathbf{r})$ can, however, still be defined as in (5). The magnitude of the imaginary part of $\mathbf{K}(\mathbf{r})$ (generalized conductivity) is larger than that of the real part (generalized susceptibility) by a factor of order β and is of the order of between 0 and β^{-1} times the expression (10).

§ 4. ELECTRODE CURRENTS

In the preceding sections we have derived formal expressions for the space currents $\mathbf{I}(\mathbf{r}, t)$, and hence for the electric displacement $\mathbf{D}(\mathbf{r}, t)$ and complex vector dielectric constant $\mathbf{K}(\mathbf{r})$, at any point in an electron gas produced by a potential difference $Ae^{j\omega t}$ between two electrodes immersed in the gas. These expressions involve the velocity distribution of the electrons $n(\mathbf{r}, \mathbf{v})$ which we consider as given, and also $\Xi(\mathbf{r})$, a measure of the electric field at every point $\mathbf{r}$ in the medium. However, $\Xi(\mathbf{r})$ can be evaluated rigorously only if, in addition to the electrode configuration, $\mathbf{K}(\mathbf{r})$ is known throughout the medium, the relations determining $\Xi(\mathbf{r})$ being

$$\operatorname{div} \mathbf{D}(\mathbf{r}, t) = 0, \qquad \ldots\ldots(13)$$

$$\int_1^2 \Xi(\mathbf{r}) \,.\, d\mathbf{s} = 1, \qquad \ldots\ldots(14)$$

where $\int_1^2$ is any line integral from the first to the second electrode. A method of successive approximation might be used in most practical cases as follows: Evaluate $\Xi_0(\mathbf{r})$, the value obtained for $\Xi(\mathbf{r})$ when the electron density is zero everywhere; $\Xi_0(\mathbf{r})$ is real everywhere. Next evaluate $\mathbf{I}(\mathbf{r}, t)$ and hence $\mathbf{K}(\mathbf{r})$ by substituting $\Xi_0(\mathbf{r})$ for $\Xi(\mathbf{r})$ in (4) throughout the medium. Substituting this approximation for $\mathbf{K}(\mathbf{r})$ in (13) evaluate the next approximation for $\Xi(\mathbf{r})$, which can then be substituted in (4) to give a better approximation for $\mathbf{K}(\mathbf{r})$, and so on. It will be seen from (13) that if the argument of the complex $\mathbf{K}(\mathbf{r})$ does not vary much throughout the medium the argument of $\Xi(\mathbf{r})$ will also vary little and from (14) that the line integral $\int_1^2$ of the imaginary part of $\Xi(\mathbf{r})$ is zero. In the case of $\beta \ll 1$ we have seen that the argument of $\mathbf{K}(\mathbf{r})$ is small compared with $\pi/2$ everywhere and hence the argument of $\Xi(\mathbf{r})$ will be negligible throughout. For $\beta \gg 1$ the argument of $\Xi(\mathbf{r})$ depends more critically on $n(\mathbf{r}, \mathbf{v})$ and the electrode configuration, but when $n(\mathbf{r}, \mathbf{v})$ does not vary much with $\mathbf{r}$ it will in general also be small.

If the transmission line leading to the pair of electrodes is adjusted so that radiation is reduced to a minimum then the currents to the two electrodes will be, as nearly as possible, equal and opposite. Let I_e be the difference of this current (flowing to either electrode) and that flowing when no electrons are present. Then I_e is the electrode current induced by the motion of the electrons throughout the ionized gas as expressed in (4). In the appendix we derive expressions for such induced currents under quasi-stationary conditions by considering the rate of change of charge induced on a conductor by a moving point charge. Using equation (vi) of the appendix we have

$$I_e = \int \Xi(\mathbf{r}) \,.\, \mathbf{I}(\mathbf{r}, t)\, dr^{(3)}, \qquad \ldots\ldots(15)$$

where the integration is taken over the whole medium not occupied by the electrodes and $\mathbf{I}(\mathbf{r}, t)$ is given by (4).

§ 5. CONCLUSION

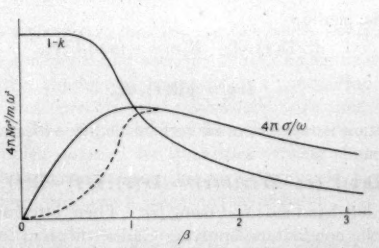
We have shown that only for the limiting case $\beta = 0$ does the electron gas behave like an isotropic, loss-free dielectric of dielectric constant K given by (10). Then $\Xi(\mathbf{r})$ is wholly real and the electrode current given by (15) is exactly in quadrature

with the potential difference. In any but this limiting case exact calculation of the electrode currents appears quite impracticable but certain qualitative conclusions can be drawn.

As β increases from zero towards unity (ω being kept constant) the component of the electrode current in quadrature (polarization current) differs from the value for $\beta=0$ by a factor of order of magnitude $1-\beta^2$ and an in-phase component (conduction current) of electrode current flows of magnitude between zero and β times the quadrature component. For small β the contributions to the in-phase current of electrons with equal and opposite velocity are nearly equal and opposite. Thus if the velocity distribution of the electrons is symmetric about zero velocity the in-phase current increases slowly at first as β increases from zero. This is indicated by the dotted line in the figure. For $\beta\sim 1$ the component in quadrature decreases more rapidly with increasing β and the in-phase component has its maximum value. For $\beta\gg 1$ the components in phase and in quadrature are of order of magnitude of or less than β^{-1} and β^{-2} respectively times the current for $\beta=0$. The electrode currents are thus of the same form as those for electrodes immersed in a medium of average dielectric constant k and conductivity σ given by

$$k=1-\frac{4\pi Ne^2}{m\omega^2(1+\beta_1^2)}; \qquad \sigma=\frac{Ne^2\beta_2}{m\omega(1+\beta_2^2)}, \qquad \ldots\ldots(16)$$

where N is the average total number of free electrons per unit space volume in the gas and β_1 and β_2 are of the same order as or smaller than β for the mean square thermal velocity of the electrons. The actual values of β_1 and β_2 depend more critically on the electrode configuration and the electron velocity distribution. A rough sketch of k and σ plotted against β (ω being kept constant) is shown in the figure.



Variation of the conductivity σ (dotted curve for symmetrical velocity distribution) and of the dielectric constant k with $\beta=v/\omega d$, ω being constant.

Expressions for the equivalent dielectric constant and conductivity of an electron gas of mathematical form similar to that of (16) were derived by Salpeter (1913). There the transit time τ of the electrons was neglected but collisions between the electrons were taken into consideration and β_1 and β_2 are replaced by the ratio of the period of the electric field to the mean free time between collisions. The physical significance of the in-phase current in both cases is that the electrons acquire energy from the electric field and dissipate part of this energy, in the one case by collisions with other electrons, in the other case by carrying energy to parts of space distant from the field. The energy exchange between electrons

and an electromagnetic field taking the finite transit time into consideration (without neglecting retardation) were calculated recently by Smith (1946) by means of quantum mechanics with special reference to electrons passing through a resonating cavity where transit time effects should be quite marked. In that paper the quantum effects are shown to become important only at very high frequencies.

ACKNOWLEDGMENTS

The authors wish to thank Mr. P. C. Thonemann and Dr. L. G. H. Huxley for helpful discussions and one of us (E. E. S.) to thank the Commissioners for the Exhibition of 1851 for a Research Scholarship.

REFERENCES

HUXLEY, L. G. H., 1947, *Wave Guides* (Cambridge: University Press), p. 297.
SALPETER, J., 1913, *Phys. Z.*, **14**, 201.
SHOCKLEY, W., 1938, *J. Appl. Phys.*, **9**, 635.
SMITH, L. P., 1946, *Phys. Rev.*, **69**, 195.

APPENDIX

(a) *A Reciprocation Theorem*

We prove here an extension of Green's reciprocation theorem in electrostatics. Consider s discrete conductors immersed in a non-conducting medium for which the dielectric displacement $\mathbf{D}(\mathbf{r})$ at any point $\mathbf{r}$ is an arbitrary function of the electric field $\mathbf{E}(\mathbf{r})$. Let charges $q_1, \ldots . q_s$ on the s conductors produce potentials $V_1, \ldots . V_s$ on the same conductors, also electric displacement $\mathbf{D}(\mathbf{r})$ and field $\mathbf{E}(\mathbf{r})$ at the point $\mathbf{r}$ in the medium; let charges $q_1', \ldots . q_s'$ produce $V_1', \ldots . V_s'$; $\mathbf{D}'(\mathbf{r})$, $\mathbf{E}'(\mathbf{r})$.

Throughout the medium

$$\operatorname{div}\mathbf{D}(\mathbf{r})=0, \quad \mathbf{E}(\mathbf{r})=-\operatorname{grad}V(\mathbf{r}), \qquad \ldots\ldots(\text{i})$$

and

$$4\pi q_j=\oint\mathbf{D}(\mathbf{r})\,.\,d\mathbf{f}_j, \qquad \ldots\ldots(\text{ii})$$

where the integration is taken over all surface elements $d\mathbf{f}_j$ on the jth conductor. We get the relation

$$\operatorname{div}\{\mathbf{D}(\mathbf{r})V'(\mathbf{r})-\mathbf{D}'(\mathbf{r})V(\mathbf{r})\}=\{\mathbf{D}'(\mathbf{r})\,.\,\mathbf{E}(\mathbf{r})-\mathbf{D}(\mathbf{r})\,.\,\mathbf{E}'(\mathbf{r})\}$$

on expanding the left-hand side and using (i). Then, integrating over the whole medium outside the conductors, applying Gauss' theorem to the left-hand side and using (ii), we have

$$4\pi\sum_{j=1}^{s}\{q_jV_j'-q_j'V_j\}=\int\{\mathbf{D}'(\mathbf{r})\,.\,\mathbf{E}(\mathbf{r})-\mathbf{D}(\mathbf{r})\,.\,\mathbf{E}'(\mathbf{r})\}dr^{(3)}, \qquad \ldots\ldots(\text{iii})$$

which is the required theorem. This theorem bears some resemblance to one enunciated by Lorentz (cf. Huxley 1947).

For many systems the right-hand side of (iii) vanishes, in which case the system is conservative. This is the case, for example, for any distribution of conductors immersed in a medium not necessarily isotropic or homogeneous but such that, in tensor notation, $\mathbf{D}_\mu(\mathbf{r})=K_{\mu\nu}(\mathbf{r})E_\nu(\mathbf{r})$ where $K_{\mu\nu}(\mathbf{r})$ is a symmetric tensor for all $\mathbf{r}$. In the more general case where $\mathbf{D}(\mathbf{r})=\mathbf{f}(\mathbf{E}/E,\ \mathbf{r})E(\mathbf{r})$ where $\mathbf{f}$ is an arbitrary vector function then the right-hand side of (iii) will still vanish for any two charge distributions which make $\mathbf{E}(\mathbf{r})$ and $\mathbf{E}'(\mathbf{r})$ parallel throughout the medium.

(*b*) *A Theorem on Induced Currents*

We prove here an extension of a theorem given by Shockley (1938). Consider three conductors (the third one being a point charge at the point $\mathbf{r}$) immersed in a medium such that the right-hand side of (iii) vanishes for the pair of distributions given by (*a*): $q_1+q_2=0, q_3=0, V_1-V_2=1$ and (*b*): $q_1'-q_2'=0, q_3'=q, V_1'-V_2'=0$. For these distributions we find the following relation between V_3 and q_1' from (iii)

$$q_1' = -V_3 q. \qquad \ldots\ldots\text{(iv)}$$

We shall suppose that the two finite conductors are "balanced" throughout all measurements such that the charges on them are always equal and opposite and that retardation of the field variables can be neglected. Let $\Xi(\mathbf{r})$ be the field produced at $\mathbf{r}$ by unit potential difference between the two conductors in the absence of any charge at $\mathbf{r}$ and let I_1 be the current flowing to the first electrode induced by the motion of a charge q at the point $\mathbf{r}$ with given velocity $\mathbf{v}$ while the potential difference between the two conductors is kept fixed. Using (iv) we have

$$I_1 = \frac{dq_1'}{dt} = -q \operatorname{grad} V_3(\mathbf{r}) \,.\, \mathbf{v} = q\Xi(\mathbf{r}) \,.\, \mathbf{v}.$$

This equation can be generalized to the case of any space current of density $\mathbf{I}(\mathbf{r}, t)$ taking the place of the single point charge and gives

$$I_1 = \int \Xi(\mathbf{r}) \,.\, \mathbf{I}(\mathbf{r}, t) dr^{(3)}, \qquad \ldots\ldots\text{(v)}$$

where the integration is taken over all parts of the medium not occupied by the two conductors.

(*c*) *Extension to the Case of a Conducting Medium*

The results of the preceding sections of this appendix can be generalized to the case of perfect conductors immersed in a conducting medium provided that all potentials and currents vary sinusoidally with time with fixed angular frequency ω and that the dimensions of the conductors are much less than c/ω. We again use complex notation for quantities varying sinusoidally with time. We restrict ourselves to media such that an electric field $\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0(\mathbf{r})e^{j\omega t}$ at a point $\mathbf{r}$ in the medium produces a space current of form

$$\mathbf{I}(\mathbf{r}, t) = \frac{j\omega}{4\pi}\left\{\mathbf{K}(\mathbf{r}, \mathbf{E}/E) - \frac{\mathbf{E}(\mathbf{r}, t)}{E(\mathbf{r}, t)}\right\} E(\mathbf{r}, t),$$

where $\mathbf{K}$ is in general complex. We define the dielectric displacement $\mathbf{D}(\mathbf{r}, t)$ as in equation (5) of the main paper. If $\rho(\mathbf{r}, t)$ is the space charge density at $\mathbf{r}$ we have

$$d\{\rho(\mathbf{r}, t)\}/dt + \operatorname{div} \mathbf{I}(\mathbf{r}, t) = 0; \quad \operatorname{div} \mathbf{E}(\mathbf{r}, t) = 4\pi\rho(\mathbf{r}, t).$$

Using these two equations and the definition of $\mathbf{D}(\mathbf{r}, t)$ we again have $\operatorname{div} \mathbf{D}(\mathbf{r}, t) = 0$. If retardation effects can be neglected the other part of equation (i) is also satisfied and so is (ii). Hence equation (iii) holds again in this case. Let the electric field produced at $\mathbf{r}$ by a potential difference $e^{j\omega t}$ between the two finite conductors (again assumed to be "balanced") be $\Xi(\mathbf{r})e^{j\omega t}$ and let the current induced to flow to the first conductor by a current $\mathbf{i}(\mathbf{r})e^{j\omega t}$ flowing at $\mathbf{r}$

be $I_1 e^{j\omega t}$ when the potential difference between the two conductors is fixed. We then have just as in subsection (*b*) of this appendix

$$I_1 = \Xi(\mathbf{r}) \cdot \mathbf{i}(\mathbf{r}),$$

provided the right-hand side of (iii) vanishes for this case. I_1, Ξ and $\mathbf{i}$ are now complex, the arguments determining the phases.

If we take as our current $\mathbf{i}(\mathbf{r})e^{j\omega t}$ the space current in the medium, $\mathbf{I}(\mathbf{r}, t)$, produced by a sinusoidal potential difference between the two conductors then $\mathbf{E}(\mathbf{r})$ and $\mathbf{E}'(\mathbf{r})$ in (iii) will be parallel automatically and hence the right-hand side will in fact vanish and we have

$$I_1 e^{j\omega t} = \int \Xi(\mathbf{r}) \cdot \mathbf{I}(\mathbf{r}, t)\, d\tau^{(3)}. \qquad \ldots\ldots (vi)$$

Thermionic Emission from Oxide Coated Cathodes

By D. A. Wright

(Communication from the Research Staff of the M.O. Valve Company at the G.E.C. Research Laboratories, Wembley, England)

MS. received 5th April 1948, in amended form 18th October 1948

ABSTRACT. It is shown experimentally that with Ba/Sr oxide cathodes, the saturated emission depends on the core material in use, and there is some correlation between total coating resistance and saturated pulsed emission. The steady-state D.C. emission is lower than the pulsed emission, but the D.C. emission measured immediately on application of the anode voltage is similar to the pulsed emission. The difference is due to a decay effect with a time constant in the range 1/1,000–1/10 sec., which is similar in many respects to the decay effect studied in an earlier paper in conductivity measurements. It is shown that the application of semiconductor theory can give a consistent account of coating conductivity and emission, provided the concentration of barium in stoichiometric excess is about 3×10^{17} atoms/cm^3. This figure has some experimental support. It is shown that the presence of an interface layer of the type actually detected should cause decay phenomena with a time constant less than 1 μsec. and in most cases less than 1/10 μsec., so that the decay from pulsed to D.C. emission is not due to the capacity effects introduced by the presence of the layer.

With thoria there is little difference thermionically when studied as a coating on tantalum or on tungsten, or as a ceramic tube, and the pulsed emission is about twice the D.C. emission. Experiment and theory indicate that thoria is an excess semiconductor containing a stoichiometric excess of about 10^{18} atoms/cm^3 of thorium. There is very little sensitivity to oxygen, however, at temperatures above 1,850° K.

§1. Introduction.

The purpose of this paper is to describe experiments in which the emission from oxide cathodes under pulsed and D.C. conditions was investigated, to apply the semiconductor theory to the emission and conductivity of cathode coatings, and to consider further the part played by interface layers between the coating and the metal to which it is applied. Such layers have been discussed in previous papers (Eisenstein and Fineman 1946, Wright 1947a), where it was shown that under certain conditions the interface layer may have considerably higher resistance than that of the coating itself, and that the associated barrier layer effect then introduces rectifier characteristics at the boundary.

In the present paper a study has been made of conventional oxide cathode coatings in which barium oxide or solid solutions of barium and strontium oxides are employed, and also of coatings of thoria.

§ 2. EXPERIMENTS WITH Ba/Sr OXIDE COATINGS

The experiments were carried out using diodes with a copper anode which could be water cooled, enabling a complete investigation of D.C. emission as well as pulsed emission to be carried out without appreciable rise of anode temperature. The cathodes were formed from directly heated strip bent into a "hairpin", permitting the central portion with area 0·1 cm^2 to be coated and to be mounted at a distance of 0·7 mm. from the anode, forming a planar system. The diode characteristics were plotted on a double logarithmic scale, which assists in determining the value of the maximum space charge limited current. This is the value required when discussing "saturated" emission, and when making Richardson plots. Different core materials were investigated, and a normal sprayed coating was compared with one in which coating was embedded in a sintered metal powder layer. The cathodes were outgassed during pumping for two minutes at 1,000° C., and barium getter was dispersed after pumping. The activation was studied under D.C. conditions until a steady state was reached, which usually required several hours' operation at 900° C. and several further hours at 820° C., taking a gradually increasing current at each temperature. The emission was then studied, to the highest possible current density in both pulsed and D.C. operation, and the causes of limitation were noted. The standard pulse length was 2 μsec., though this could be varied from $\frac{1}{2}$ to 35 μsec. when required. The repetition rate was maintained at 50 c/s. The more interesting cathodes were operated for long periods to determine the changes in performance with time.

Some typical results with Ba/Sr oxide coatings are shown in Figure 1 in which the full lines refer to pulsed emission. In all cases the current, after deviation from the space charge law, continues to increase with voltage, rapidly at the higher temperatures, until it is limited by the passage of discharges between cathode and anode. This phenomenon has been referred to previously by the writer (Wright 1947a), and is described by Coomes (1946) as cathode sparking.

All the curves in Figure 1 for temperatures above 700° C. are continued beyond saturation to the point at which persistent sparking prevented further increase in voltage and current.

The curves for O–nickel are often, but not always inflected as drawn in the saturated region. When not inflected, they are of the form shown for pure nickel. The curves for nickel with 0·1% Al and for pure nickel are less frequently inflected, and are usually as drawn. With pure nickel it is sometimes observed that the curves are inflected during the activation process, but become linear when activation is complete. The writer (1947a) suggested that such inflected characteristics might be associated with the interface, since the resemblance to a rectifier characteristic in the "difficult" direction of flow is clear. There are considerable difficulties associated with this theory, which is by no means proven, but it will be observed that the voltage drop which must be supposed to occur across the interface on O–nickel, both when the upward inflection occurs and when sparking begins, is consistent with this view, when the thickness is 10^{-4} cm. and the resistance is some tens of ohms per cm^2.

Details concerning the interfaces on the different core materials are given in Table 1. Here the identification of the interface compound and the estimation of its thickness were carried out by Rooksby (1947). The resistivity of barium orthosilicate is known to be high (Eisenstein 1947). Assuming that the observed

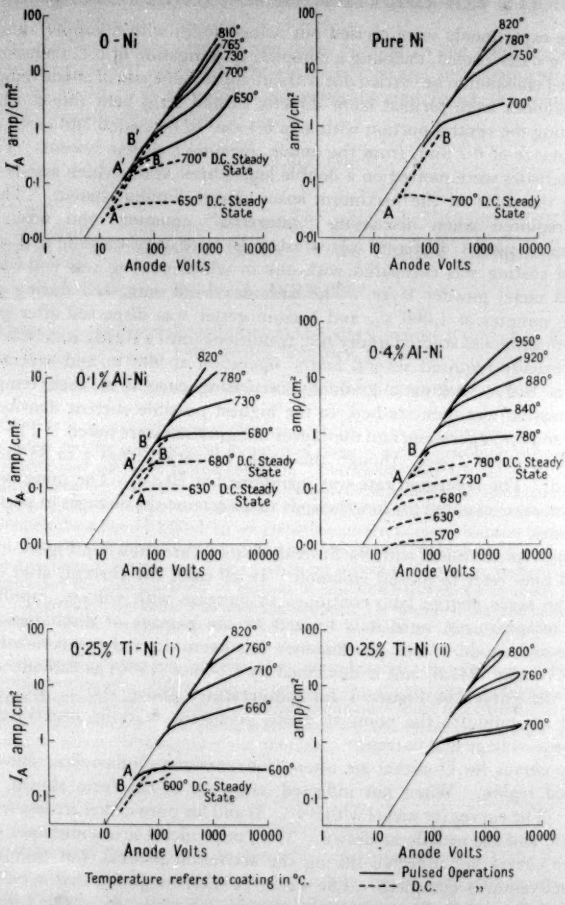


Figure 1. Diode characteristics with various core materials.

variations in coating resistance are due mainly to the resistance of the interfaces formed with the different cores, indications of interface resistances in the present work were obtained in two ways: (*a*) by observing the temperature rise of the coating when taking 1·5 amp/cm² in D.C. operation from each cathode when the temperature was initially 830° C. and (*b*) by plotting coating temperature against

Table 1. Composition and Resistance of Core–Coating Interfaces

	Pure Ni	O–Ni	0·4% Al–Ni	0·1% Al–Ni	0·4% Si–Ni	0·25% Ti–Ni	Pt
Nature of Interface	None detected	Magnesium + oxide Ba ortho-silicate Ba_2SiO_4	Barium aluminate $BaAl_2O_4$	Barium aluminate $BaAl_2O_4$	Barium ortho-silicate Ba_2SiO_4	Barium ortho-titanate Ba_2TiO_4	Barium platinate
Thickness of interface	—	10^{-4} cm.	$\sim 5\times10^{-4}$ cm.	5×10^{-5} cm.	10^{-3} cm.	10^{-3} cm.	Up to 10^{-3} cm.
Indicated resistivity of interface relative to coating	—	High	A little higher	—	High	Variable, resembles coating in results (i) higher in (ii)	High

The O-Ni contained 0·10% Mg and 0·05% Si. The specification also permits traces of Co, Mn, Fe, Cu. The remaining nickels were of high purity apart from the additives tabulated.

Table 2. Relation between Coating Resistance and Nature of Core Metal

	Pure Ni	O–Ni	0·1% Al–Ni	0·4% Al–Ni	0·4% Si–Ni	0·25% Ti–Ni (i)	0·25% Ti–Ni (ii)	Pt	Ni with 0·18% Si 0·06% Mg
(*a*) Temperature rise in D.C. operation at 1·5 amp/cm² at 830° C.	7°	15°	6°	30°	—	10°	35°	35°	30°
(*b*) Total coating resistance at 790° C. determined by temperature changes (Ω/cm²)	3	4	3·5	7	20	4	—	6	5

Coating is 10 mg/cm² of Ba and Sr carbonates in equimolecular proportions.

current drawn in D.C. operation in a diode with indirectly heated cathode where the coating surface could be observed through a hole in the anode. The temperature at first falls due to the cooling term $I\phi$ where I is the current density and ϕ is the work function in electron volts, but heating occurs due to the term I^2R, where R is coating resistance. When the temperature returns to its original zero current value, $I^2R = I\phi$, and from this value of I, R can be determined. Some values are shown in Table 2.

The dotted curves in Figure 1 indicate the behaviour of the emission in D.C. operation. It is seen that the dotted lines are lower than the pulsed curves at the corresponding temperatures. Richardson lines could be plotted confirming this difference in value. The ratio of pulsed to D.C. saturated emission lies in the range 2 to 20 though it is most commonly between 5 and 10 at 740° C. The reasons for the difference are of considerable interest, and no completely satisfactory theory has yet been produced. Two types of decay effect have been observed previously with Ba/Sr oxide cathodes, as mentioned by Coomes (1946). The first has a time constant of the order of microseconds or tens of microseconds, and the second a time constant of the order of a hundredth or a tenth of a second. The first has been encountered in the present experiments only in specific cases, as shown in Table 3, where the results of Figure 1 are shown, together with those of other core materials where the shape of the characteristics had no special interest. None of the cores in regular commercial use showed any decay of emission over the period $\frac{1}{2}$ to 35 μsec. with a double oxide coating, though a decay was encountered with single barium oxide on O–nickel. The second decay with the longer time constant did, however, occur in almost all cases, and it is this decay which caused the difference between pulsed and D.C. emission. This decay was first studied by making D.C. measurements with the cathode below 700° C. By presetting the anode voltage to the required value, and observing the variation of the current with time after switching on, the variation of the decay effect with current density was investigated. After reaching a steady state, the cathode was run at 800° C. for one minute with no current before taking further observations. It was found that at low current density there were no time effects, but as current was increased a decay effect appeared whose magnitude increased with current density while its time constant decreased. At low temperature and low current density the time constant was of the order of a second. Measurements on D.C. instruments, therefore, enabled the decay effects to be studied and the zero-time current to be estimated, i.e. the current obtained immediately on switching on the anode voltage. The decay effect appeared while the zero-time current was still space-charge limited, for example at A in Figure 1, and whereas the zero-time current continued to follow the space-charge law AB, the steady state current saturated at a value not much greater than that at A. The zero-time current could in some cases be followed until saturation occurred at the same current density as in pulsed operation, showing quite definitely that it is this decay effect which causes the difference between pulsed and steady-state D.C. emission. All these decay effects are very similar to those observed in the study of conduction current in cathode coatings (Wright 1947a) and there seems little doubt that a common origin must be sought.

The work function ϕ was determined by making Richardson plots of $\log I/T^2$ against $1/T$, where I is the maximum space charge limited current according to the formula

$$I = AT^2 \exp(-\phi/kT). \qquad \ldots\ldots(1)$$

The value for the work function ϕ for the O–Ni and Al–Ni cathodes in Figure 1 is 1·2 ev. in pulsed operation, and 0·95 in D.C. operation. The Richardson plots for pulsed and for D.C. emission are thus not parallel, and would intersect if extrapolated to low temperatures. The ratio between D.C. and pulsed emission is therefore a function of temperature. It may be noted that the results with pure nickel and with titanium nickel as given in Figure 1 indicate a higher work function, which remains however to be confirmed.

The results with Ti–Ni were rather variable, and are included here to show that there was a correlation between pulsed emission, limiting current density and coating resistance, which could also be observed comparing different core

Table 3. Behaviour of Ba/SrO Cathodes with different Core Materials in the best State of Activation

(1)	(2)		(3)		(4)	
	Pulsed	D.C.	Pulsed	D.C.	2 μsec.	35 μsec.
Pure Ni	8	0·3	60	1·2	—	—
O–Ni	6	0·9	45	1·8	—	—
Ti–Ni 0·25% (i)	12	3·0	83	3·0	—	—
Ti–Ni 0·25% (ii)	6	0·3	25	1·8	Slight	Large
Al–Ni 0·1%	7	0·6	30	1·5	—	—
Al–Ni 0·4%	0·5	0·2	1·5	1·0	—	—
0·18% Si–Ni with 0·06% Mg	4	0·6	40	1·5	—	—
Si–Ni 0·4%	1·5	0·07	15	0·3	—	—
Pt	1	0·2	12	0·8	—	Slight
W powder	4	0·6	20	1·2	—	—
Th powder	0·8	0·15	6	0·4	—	—
Be–Ni 0·5%	0·02	0·01	—	—	—	—
Ta	0·03	0·02	—	—	—	—
Fe	0·6	0·07	10	0·15	—	—
Embedded coating in :						
Ni (" Mush ")	1·0	0·04	3	0·5	Slight	Large
Tungsten	3	0·05	10	1·2	Slight	Large
BaO on O–Ni	0·4	0·06	0·6	0·15	Slight	Large

(1) Nature of core metal to which coating is applied; (2) saturated emission at 740° C. in amp/cm² under pulsed (2 μsec.) and D.C. conditions ; (3) maximum current density which can be drawn at 790° C. ; under pulsed conditions this is greater than the saturated emission and is limited by sparking ; in D.C. operation various causes limit the current density (see text) ; (4) decay effects in pulsed operation.

The coating is 10 mg/cm² of Ba and Sr carbonates in equimolecular proportions.

materials in Tables 2 and 3. The results with Ti–Ni refer (i) to the best observed, and (ii) to the worst, and it is interesting that (i) is the only case where pulsed emission and limiting current are higher than with very pure nickel, while the coating resistance is of the same order. In (ii) the results are much poorer, but the reason for this spread is not understood. In typical extreme cases (i) and (ii) the interface was of the same composition, barium orthotitanate, and of the same thickness, i.e. about 10^{-3} cm., thus either the spread is not associated at all with the interface, or the interface resistivity can vary without detectable change in composition from a value lower than that of the coating itself to a value considerably higher.

The relation between interface thickness, coating resistance, and pulsed emission is shown most clearly by comparing the behaviour of the 0·1% and the 0·4% Al–Ni.

The cores quoted in Table 3 as tungsten and thorium powder refer to cathodes in which a layer of metal powder was sintered to a tantalum base, and the coating was sprayed in the usual way on to the sintered layer. In the "embedded" coating, a layer of metal powder was again formed by sintering, but with larger particle size and greater thickness than in the previous case, to form a porous structure. The coating was painted into the layer, no excess surface coating being permitted. This procedure with nickel powder sintered to a nickel strip corresponds with the "mush" cathode used in the cavity magnetron.

In all cases the current drawn under pulsed conditions could be increased above its "saturated" value by increase of anode voltage, until finally it was limited by sparking, as in Figure 1. In D.C. operation however at 790° C. the continued increase in voltage and current density was limited by various factors. Sometimes there was limitation due to the formation of incandescent spots on anode or cathode, or to the temperature rise of the coating becoming uncontrollable. This might occur either below or beyond saturation. "Poisoning" of the emission due to gas from the anode was a limitation in some cases. Thus the D.C. figures in column 3 in Table 3 are not necessarily typical of the cathodes, unlike the figures in the other columns.

In continued D.C. operation at high current density, the emission fell and became the limiting factor in most cases after some tens of hours' operation at 790° C. taking more than 1 amp/cm² initially. Thus although current density quoted in Table 3 exceeds 1 amp/cm² in many cases in D.C. operation, these high figures were not maintained in an extended life test.

Some points of interest arise in connection with activation. As the diodes were sealed off from the pump immediately after outgassing the cathode, the emission was poor in all cases at the beginning of operation. It was best with O–Ni, approximately ½ amp/cm² at 900° C., about half this with 0·1% Al–Ni, but much lower in all other cases. With O–Ni and Al–Ni, considerable activation occurred by running at 900° without taking emission, but in all other cases electrolytic activation was necessary to raise the D.C. emission above a few ma/cm² at 740° C. This activation was continued for some tens of hours in most cases before the emission developed the high values shown in Table 3. In the cases of O–Ni and Al–Ni, although there was activation by temperature alone, it was necessary to activate electrolytically to achieve the figures in Table 3, and while most of the activation occurred in less than one hour taking about 1 amp/cm² at 900° C., there was a slow increase over a long period, so that the total activation time was again some tens of hours.

Since the pulse emission is larger than the steady-state D.C. emission, it is of interest to consider the behaviour when alternating voltage is applied between anode and cathode. Some brief tests were carried out in which a sine wave voltage was applied at 50 and at 500 c/s. The current was measured using an oscillograph, and again at the lower cathode temperatures, i.e. below 750° C., it was possible to observe that the zero time emission resembled the pulsed emission, while a rapid decay effect led to a steady state in which the saturated peak emission was intermediate between the pulsed and D.C. emissions. In cases where the

pulsed to D.C. ratio was about 10, the A.C. to D.C. ratio was about 3. Similarly in such cases the limiting current density in A.C. operation was about three times as large as in D.C. operation.

§3. APPLICATION OF SEMICONDUCTOR THEORY

While there is no certain evidence to show that de Boer's theory, in which he attributes the emission to thermal ionization of an adsorbed layer of barium atoms, (de Boer 1935) is incorrect, it is of interest to develop the semiconductor theory, which has been considered in many later writings, for example, Blewett 1946, Coomes 1946, Nishibori and Kawamura 1940, Morgulis 1947. According to this view the coating is a semiconductor in which a stoichiometric excess of barium provides the activating "impurity". Whether this occurs due to the formation of interstitial barium atoms (Frenkel defects) or to the development of an oxygen ion deficiency (Schottky defects) is not clear, although the former is improbable in the sodium chloride type of lattice (Seitz 1940). According to this theory, the emission is one of electrons from the conduction band of the semiconductor, the work function being determined by the barium concentration in the coating and on its surface. The surface effect corresponds closely to that of barium, thorium or caesium in lowering the work function of tungsten. The writer has shown (Wright 1948) that when the coating is fully activated from the point of view of internal barium concentration, as estimated from the conductivity, the work function at room temperature should not exceed 1·6 ev. in the absence of surface barium, and is probably less. Thus the effect of surface barium if present is to lower the work function by at most 0·5 ev. This, of course, can have a considerable effect on emission, and it is not intended to deny any importance to surface barium, but rather to emphasize the importance of the internal concentration.

We proceed to take the observed value of the coating conductivity after activation, and from the semiconductor formulae to calculate the excess barium concentration. Then from this concentration and the observed work function we calculate the thermionic emission, and compare with observed values.

In order to calculate the emission and conductivity, the coating may be treated as a semiconductor with energy levels as in Figure 2 (p. 199). The bottom of the conduction band is situated χ ev. below the zero level, and the "impurity", i.e. the excess barium, has an electron energy level at depth 2ϵ below the bottom of the conduction band. When the number of electrons in the conduction band is small compared with the number of excess barium atoms, the "activation energy" for conductivity has the value ϵ, and the number of electrons per cm^3 in the conduction band is given (Fowler 1936) by:

$$n = n_0^{\frac{1}{2}}\{(2\pi \boldsymbol{mk}T)^{\frac{3}{2}}/\boldsymbol{h}^3\}\exp(-\epsilon/\boldsymbol{k}T) \qquad \ldots\ldots(2)$$

where n_0 is the number of excess Ba atoms per cm^3, $\boldsymbol{m}$ is the mass of the electron, T is the absolute temperature, $\boldsymbol{k}$ is the Boltzmann constant and $\boldsymbol{h}$ is Planck's constant.

In order that an electron shall be emitted thermionically from the semiconductor, it must, after gaining the energy ϵ, gain the further energy χ, and it

is therefore possible to define a work function $\phi_1 = \chi + \epsilon$ which appears in the thermionic emission formula for semiconductors:

$$I = DA_0 T^2 n_0^{\frac{1}{2}} \frac{\boldsymbol{h}^{\frac{3}{2}}}{(2\pi \boldsymbol{mk} T)^{\frac{1}{2}}} \exp\left(\frac{-(\chi+\epsilon)}{\boldsymbol{k}T}\right). \qquad \ldots\ldots(3)$$

Here A_0 is the usual thermionic constant $4\pi \boldsymbol{emk}^2/\boldsymbol{h}^3$, where $\boldsymbol{e}$ is the electronic charge. A_0 has the value 120 amp/cm^2/degree^2. D is a term introduced to take account of reflection at the surface potential barrier of a solid. With clean metals however D is nearly unity, and the fact that observed values of A in equation (1) are less than A_0 is due in the case of metals to variation in ϕ with temperature, and not to a small value of D. With adsorbed films however D may become small, as with thoriated tungsten. With oxide cathodes D may be small if there is a surface layer of adsorbed barium analogous with thorium on tungsten, but in the absence of such a layer, the rather large expansion coefficient (27×10^{-6}, Eisenstein 1946) will lead to a large variation in ϕ with temperature and therefore a value of A much smaller than A_0.

It is important to notice that while equation (3) is the correct formula for emission from a semiconductor, in practice in the oxide cathode the coating is in contact with the metal, and there is in equilibrium an adjustment of energy levels such that the bottom of the conduction band is at a height U above the top of the Fermi distribution in the metal, where U is given by:

$$U = \epsilon - \boldsymbol{k}T \ln\{n_0^{\frac{1}{2}} \boldsymbol{h}^{\frac{3}{2}}/(2\pi \boldsymbol{mk} T)^{\frac{1}{2}}\}. \qquad \ldots\ldots(4)$$

The emission formula (3) can therefore take the alternative form:

$$I = DA_0 T^2 \exp\{-(\chi + U)/\boldsymbol{k}T\} \qquad \ldots\ldots(5)$$

in which $\chi + U$ is the effective work function of the system related to the metal. When Richardson plots are made therefore according to equation (1), the slope ϕ is equal to $\chi + U$. Since, however, U itself varies with temperature, this procedure is not strictly correct, and the proper procedure with a semiconducting coating is to plot log $I/T^{5.4}$ against $1/T$, as is evident from equation (3). When this is done with the results in Figure 1, the value of ϕ, for coatings on O–Ni and on Al–Ni becomes 1·1 ev. and this is therefore the value of $(\chi + \epsilon)$. Strictly this is the value of ϕ_1 at the absolute zero of temperature.

The conductivity of the semiconductor is given by

$$\sigma = n\boldsymbol{e}v \qquad \ldots\ldots(6)$$

where v is the electron mobility in cm/sec. per unit voltage gradient. The value of v is not known for the oxide coating, and is not readily predicted. In ZnO it is known to be about 10, and this value will be assumed for the moment to apply here also. The conductivity at 1,000° K. is about 3×10^{-3} ohm^{-1} cm^{-1} as shown by the above results, and by Eisenstein and Fineman (1946). The value of n at 1,000° K. is therefore 3×10^{15} with $v = 10$.

It is known in general that the value of ϵ is dependent on that of n_0, decreasing as n_0 rises, so that with increase in n_0 the value of n increases rapidly. There have been very variable results in the past in experimental values of both n_0 and ϵ, and often they have not been related with the thermionic state of the cathode. For the present cathodes when well activated, ϵ has the value 0·7 ev. as found from probe measurements of conductivity. With this value equation (2) indicates that with $n = 3 \times 10^{15}$ at 1,000° K., n_0 is about 2×10^{17}. This value is lower than many earlier experimental estimates, but is supported by measurements at these

Laboratories so far unpublished. Dr. R. O. Jenkins has made measurements in which an activated cathode was released into a separate container which was "sealed off", i.e. isolated completely from the diode, and connected to apparatus for micro gas analysis. The production of hydrogen by reaction between water vapour and free barium gave a measure of the free barium in the coating. The values obtained were about 3×10^{17} atoms/cm^3. Similar values have been obtained in measurements made at the Bell Telephone Laboratories by the same method, according to results communicated privately to the author by Dr. White and Dr. Wooten.

With n_0 equal to 3×10^{17} equation (3) gives with $\chi+\epsilon=1{\cdot}1$ and D equal to unity, an emission of 40 amp/cm^2 at 1,000° K. The observed pulsed emission (Table 3) is of the order one-tenth of this, and as noted above a value of D of one-tenth could occur as a result of either an adsorbed layer of Ba or change in ϕ_1 with temperature. The semiconductor model developed above is therefore able to describe consistently both the conductivity and the pulsed emission. The steady state D.C. emission is lower as discussed above, and the decay of emission with time leading to this state is considered further in the following section.

§ 4. DECAY EFFECTS

Several mechanisms exist which could lead to decay effects, and it is of interest to list them. Regarding the coating as a purely electronic conductor in good contact with the core metal, in equilibrium conditions there is a space charge at both boundaries, at the outer one due to surface states (Bardeen 1947), and at the inner one due to a Schottky depletion or exhaustion layer. When current flows these conditions are disturbed, and time effects must occur, though there is no evidence yet of the time constants to be expected. Another mechanism causing decay is provided if the coating is separated from the core by a high resistance interface, since this layer will behave as a leaky condenser. The time constants for these effects are considered below.

When electrolytic flow is considered, other possibilities appear, for example, a simple polarization, or the change in concentration of the appropriate ion at either or both of the boundaries may directly influence the flow of electrons across the boundary. A direct effect of this type at the outer boundary was considered by Blewett (1939) and by Sproull (1945). An effect at the inner boundary has been postulated by Biguenet (1947), who proposes that positive ions of barium absorbed at the inner (core–coating) boundary lead to a high zero-time emission, but the electron flow neutralizes the ions causing decay of the flow across the boundary.

Evidence for associating the decay with phenomena at the inner boundary was indicated by the writer (1947a). All these effects are associated with the two coating boundaries, but in addition it seems possible that both space-charge and ion concentration effects might occur at the boundaries of individual coating crystals, which would lead to a dependence of time effects on crystal size. The accumulation of negative ions within the coating at the positive (vacuum) boundary, or at the positive boundary of individual crystals, would lead to decay of both conduction and emission current.

Detailed consideration of the effect of this accumulation on the energy levels shows that it might account for the observed decrease in work function during decay. This will be dealt with elsewhere. It involves quite a different picture

of the detailed process from that of Sproull, which would lead to an increase in work function. Sproull actually observed a decrease in work function during short period decay, as in the present case during the longer period decay.

In addition to these problematical effects, there are two causes of time changes which may be practically of greatest importance. Firstly, there is the cooling of the coating due to electron evaporation, $I\phi$, and the heating I^2R already referred to, and secondly the "poisoning" of the emission by gas from the anode or glass envelope.

Temperature changes due to $I\phi$ cooling and I^2R heating were studied as above in connection with estimation of resistance, and it was noted that in no case was the temperature drop, due to $I\phi$ cooling, greater than 15° C. The fall in emission during decay by a factor of 5 to 10 would require a temperature drop of the order 40°, thus this cooling effect could not explain the observed decay. The poisoning due to gas from the anode is a possibility which cannot be ruled out in the present case, since although the anode was maintained cool, gas liberation under electron bombardment is known to occur. Two facts suggest that the decay process is at least in part a fundamental property of the oxide coating, not due to extraneous gas: (i) the fact that after decay the work function is less than at zero time and (ii) the fact that decay occurred during conduction current measurements, when no current flow occurred to the anode (Wright 1947a). It will be seen from the following section that the capacity of the interface layer does not cause decay effects with the observed time constant, thus the observed decay is due either to electrolytic effects, or to "poisoning". Further experiments are required to show which of these is the most important with a fully activated cathode.

§ 5. EFFECT OF A HIGH-RESISTANCE INTERFACE

The relation between current flow and voltage drop across a barrier layer is of the type sketched in Figure 3, if we consider the difficult direction of flow. This is when electrons flow from metal to semiconductor, when the semiconductor is of the excess type. The whole system can be considered in an elementary manner as in Figure 4, if we recall that R_1 varies with current density as in Figure 3, and that R_3 varies according to the law for current in a diode.

If a voltage E_A is applied between cathode and anode at time 0, we find in this system at time t

$$i_1 = \frac{E_A}{R_1+R_2+R_3}\left[1-\exp\left(-\frac{R_1+R_2+R_3}{CR_1(R_2+R_3)}t\right)\right],$$

$$i_2 = \frac{E_A}{R_2+R_3}\exp\left(-\frac{R+R_2+R_3}{CR_1(R_2+R_3)}t\right).$$

Thus, at time 0, $i_1=0$ and $i_2=E_A/(R_2+R_3)$. In the steady state $i_2=0$ and $i_1=E_A/(R_1+R_2+R_3)$.

Thus no decay effects will be detectable in a high impedance diode where R_3 is always large compared with R_1 and R_2, nor in any diode when the voltage E_A is sufficient to draw saturated emission, since R_3 is then again large. Decay effects will be observed, however, with space-charge limited currents if R_1 is appreciable compared with R_2+R_3. The time constant of the decay is $CR_1(R_2+R_3)/(R_1+R_2+R_3)$ which of course is always less than CR_1. It is thus possible to obtain characteristics as in Figure 5. Here the dotted curve is the current at zero time $i_2=E_A/(R_2+R_3)$ and the full curve is the steady state

current $i_1 = E_A/(R_1+R_2+R_3)$. This is obtained by modifying the characteristic in the space-charge limited region by the addition of a resistance R_1, which varies as in Figure 3.

If the dielectric constant of the interface layer is of the order of 10, the capacity per cm^2, when the thickness is 10^{-4} cm. will be of the order of 0·01 μF. Thus with O–nickel we may expect a time constant CR_1 of the order 0·1 μsec. A thicker layer with similar resistivity and dielectric constant will have a similar time constant. As the time constant for the system is, as we have seen, less than CR_1, when round-topped pulses of about 2 μsec. are studied, we may expect only to observe the full curve of Figure 5, i.e. the steady state. A greater time constant will only be encountered if the resistivity or dielectric constant are higher than the values we have assumed. It is clearly possible that cases may occur where the time constant becomes as large as 1 μsec., but greater values

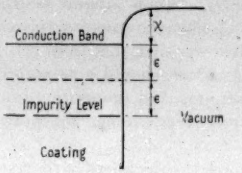


Figure 2. Energy levels in coating.

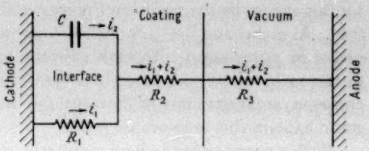


Figure 4. Representation of diode with coating and interface.

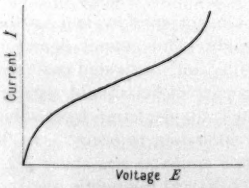


Figure 3. Current-voltage plot for a barrier layer in the difficult direction of flow of electrons.

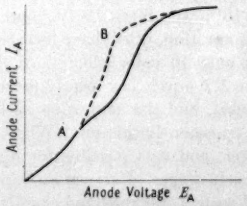


Figure 5. Diode characteristic modified by interface resistance.

are improbable. The writer thought it possible at one time that this decay due to the capacity and resistance of the interface layer could explain the difference between pulsed and D.C. emission. However, as the present results indicate a time constant of the order of 0·01 second, this is not possible, and it seems now that one of the other mechanisms discussed above is the cause of this slower decay. This is confirmed since, in the earlier stages of activation looped characteristics as in Figure 5 are frequently encountered, but the D.C. emission is then considerably less than the value at the lower "knee" A.

Thus if an interface layer is present, it will cause in well activated cathodes decay effects with such short time constants that its effect will be fully established in pulsed operation with 2 μsec. pulses, and the possibility is confirmed which is referred to in § 1, that the interface layer may be the cause both of the inflected characteristics of Figure 1 and of the sparking which finally limits current density

in pulsed operation. It was established (Wright 1947a) that when a winding of nickel with 0·4% Al was used as the cathode, the emission and conduction current showed a saturation effect at the same current density, suggesting that the core–coating boundary was controlling current flow in both cases. The increase in current with voltage beyond the "knee" was rapid in both cases and it was not certain whether this saturation effect was one which under normal diode conditions would correspond with emission saturation, or only with minor departures from the 3/2 power law at current densities below saturation. The former assumption was made when attributing the inflected characteristic to the interface and when suggesting that when a high resistance interface is present the emission may be determined at the inner boundary. This was the opinion of Lowry (1930) and initially of Reimann (1934). According to this view, when the interface is absent the emission is determined at the outer boundary, as occurred in Becker's experiments (Becker 1929). Some support is offered for this theory by the results in Figure 1 and Table 3, since the emission from both 0·4% Al nickel and 0·4% Si nickel is considerably lower than that from 0·1% Al nickel or pure nickel. In each case the reducing action should give according to the earlier views a greater free barium concentration and therefore a greater emission, and restriction of flow through the interface layer is a mechanism which could explain this unexpected result.

With a core material where no interface is formed, the thickness of the barrier layer formed by the loss of electrons to the metal (Schottky depletion layer) is given by $t=\sqrt{(\kappa\phi/2\pi n_0 e^2)}$ where κ is the dielectric constant of the semiconductor, and ϕ the height of the potential barrier at the boundary. With $\kappa=10$, $\phi=1$ ev., $n_0=3\times10^{17}$, we obtain $t=10^{-5}$ cm. While this is not thin enough for tunnel effect penetration, it will have little restrictive effect on current flow, and a voltage drop of only 10 volts will reduce the barrier height. Only minor departures from the 3/2 power law should therefore occur with a fully activated coating on pure nickel, and the saturation in Figure 1 with pure nickel should represent coating emission limitation. When partly activated, the depletion layer should be thicker, and it is possible for inner boundary limitation to occur. In fact in the early stages of activation when the saturated emission was about 1/3 of that in Figure 1, the inflected characteristic as for O–Ni was frequently observed. In result (i) with titanium nickel, the behaviour of the coating resistance (Table 2) and the emission (Figure 1) indicates that the layer of barium orthotitanate had a resistivity of the same order as that of the coating itself, unlike the orthosilicate and the aluminate. In result (ii) however, the resistance was considerably higher. With Pt the layer of barium platinate has apparently a resistivity higher than that of the coating. Thus the conclusions concerning the possible effects of the interface layer are:—If the layer has high resistivity and is sufficiently thick, it may cause a limitation of current flow across the core–coating boundary, which may appear as a limitation of emission from the cathode. Such a limitation if it occurs will already exist when taking pulsed emission in microsecond pulses, and will not be the cause of the decay effects with time constants longer than 10 μsec. This type of limitation will probably lead with thin layers to the inflected type of characteristic in which, after the inflection, the current rises with voltage almost as steeply as before saturation. These effects are not encountered with pure Ni, and rarely with 0·1% Al–Ni or Ti–Ni, when fully activated, but may be the cause of the behaviour of nickel containing more than 0·1% of Si,

of O–Ni and of 0·4% Al–Ni. Thus the departures from Ohm's law and other effects, discussed by the writer (1947a) are to be attributed to the core materials used, i.e., O–Ni containing Mg and Si as the core proper, and Ni with 0·4% Al as the probe winding, and would not be expected with purer nickels. The dependence of the decay phenomena on the core material is confirmed, since for example, pure Ni and Ti–Ni (i) with similar pulsed emission, have very different D.C. emission.

§ 6. EXPERIMENTS WITH THORIA

Experiments of the same type as those described in § 1 were carried out with thoria coatings sprayed on to a directly heated tantalum strip. The results have already been briefly reported (Wright 1947b), and here some features of these results will be amplified a little, while further experimental results are described. With tungsten strip the adhesion was so poor that few results could be obtained, but with tantalum, while the adhesion was not reliable, good cathodes could be made and tested. The Richardson plots were made according to the equation (1), taking the maximum space charge limited current as the saturated emission. The results of D.C. tests gave for A a mean value of 2·5 and for ϕ a value of $2{\cdot}5_4$ ev. The work function in pulsed emission was slightly higher, $2{\cdot}6_2$ ev., and the value of A was 7·5 corresponding with an emission a little more than twice as great. During continued D.C. operation a decrease in emission took place when taking currents of the same order as the saturated emission, i.e. 3–10 amp/cm^2 at 1,950–2,050° K., at a rate which depended in this range more on temperature than on current density. Tests have not been made taking currents small compared with the total emission, but stability would be expected over a long period when drawing 0·5 amp/cm^2 or less, as was found in Weinreich's experiments (1945). There was little change in the appearance of the coating in 500 hours' operation, though slight darkening may have occurred.

With thoria painted into a layer of sintered tungsten powder on tantalum, the behaviour was similar, except that the A value was lower, and that in continued operation the coating became grey until it could only with difficulty be distinguished from the tungsten. In the later experiments, a tube of thoria was made by the normal ceramics technique of extrusion and sintering at 1,800° C. The tube was of 3·5 mm. diameter and 0·5 mm. wall thickness, and was indirectly heated using an internal tungsten heater spiral which was made to slide easily into the tube, presumably establishing intermittent contact. The ends of the tube were supported in tantalum-cups, into which the heater extended, so that the thoria temperature was uniform along its length. The pulsed emission reached its maximum value within a few minutes of outgassing, giving the constants $A=3{\cdot}5$, $\phi=2{\cdot}5_5$ ev. In all cases the emission fell during continuous pulsed operation, to about 50% of its initial value in 200 hours. Slight darkening occurred in this time, and it is possible that the fall in emission both with the tubes and with the coated strip was due a change in emissivity, i.e. to drop in the true temperature corresponding with an observed pyrometer temperature. There was no increase in work function ϕ in any experiment. The cathodes were resistant to emission "poisoning" by gas, for example, both ceramic tube and coated strip cathodes have been operated in diodes without getter, and have behaved in all respects as in a well gettered diode. There is, therefore, no suggestion that the emission decay could be due to gas.

Further tests on the effect of gas have been carried out using thoria tube cathodes. Hydrogen or oxygen were introduced into contact with the cathode at a pressure of 10^{-3} mm. in a vacuum system with total volume of 1 litre. At temperatures above 1,850° K. the effect of this pressure of oxygen was to lower the emission by about 10%, while recovery was rapid on pumping out the gas. At temperatures between 1,700 and 1,800° K. the emission fell by about 40%, and below 1,500° K. fell to about one quarter of its initial value. Recovery in vacuum was slow at the lower temperatures, but rapid above 1,850° K. The effect of exposure to the atmosphere when cold was similar to that of exposure to the low pressure of oxygen at any temperature below 1,500° K. No effect due to hydrogen was detected at any temperature from room temperature to 1,950° K., provided that a hydrogen experiment was not carried out immediately after an oxygen experiment. If this was done, small decreases in emission were observed, probably due to a water vapour reaction caused by atomic hydrogen combining with traces of oxygen. Thus the emission is not affected by hydrogen, but is slightly decreased by oxygen at operating temperatures. The sensitivity to oxygen is very much less than that of thoriated tungsten, even when the latter is carbonized.

§ 7. APPLICATION OF SEMICONDUCTOR THEORY TO THORIA

The resistance measurements (Wright 1947b) show that if thoria is treated as a semiconductor, the value of ϵ in Figure 2 is 1·1 ev., while at 1,900° K. the resistivity is of the order 30 ohm-cm. If it is again assumed that the electron mobility is about 10 cm/sec. per v/cm., the application of equation (6) gives n of the order 10^{16} at 1,900° K. and n_0 therefore about 10^{18} atoms/cm^3, from equation (2). When $\log I/T^{5\cdot4}$ is plotted against $1/T$, according to equation (3), the slope gives a value for ϕ_1 of 2·50 ev. When this is substituted in (3) together with $n_0 = 10^{18}$, the value of I at 2,000° K. with $D = 1$ is 14 amp/cm^2. The observed current is 5·5 amp/cm^2; thus the results are consistent with observation provided that D is 0·4. This is larger than would correspond with an adsorbed layer of Th, but is consistent with a "clean" surface and a value of D entirely due to the effect of thermal expansion on the work function. The thermal expression of ThO_2 is 11×10^{-6} (Eisenstein 1946). With a value of n_0 of 10^{18} atoms/cm^3, it is clear that thoria is not an intrinsic semiconductor, since n_0 would then have a value of the order 10^{22}. The behaviour of the emission on exposure to oxygen suggests that thoria is an excess semiconductor, and it is natural to suppose that thorium atoms in stoichiometric excess represent the impurity. The very rapid activation is difficult to explain by this theory, but otherwise the behaviour is consistent throughout with this view. The nature of the darkening of thoria in operation comes into question here; it is probably due in the indirectly heated tube, and in the coating embedded in tungsten powder, to diffusion of tungsten through the thoria, rather than to the formation of visible free thorium. Whether this tungsten diffusion reduces the emission directly, or only apparently as a consequence of thermal emissivity changes, remains to be decided.

ACKNOWLEDGMENT

The author desires to tender his acknowledgments to the M.O. Valve Co., Ltd., on whose behalf the work described in this publication was carried out.

Note added in proof. When this paper was prepared the author did not know of the treatise *Die Oxydkathode* by G. Herrmann and S. Wagener, 1944 (Leipzig: Barth). This book includes the semiconductor theory of oxide cathodes. The theory is also treated by H. Friedenstein, S. L. Martin and G. L. Munday, 1948, *Reports on Progress in Physics*, Vol. XI (London: Physical Society), p. 298.

REFERENCES

BARDEEN, J., 1947, *Phys. Rev.*, **71,** 717.

BECKER, J. A., 1929, *Phys. Rev.*, **34,** 1323.

BIGUENET, C., 1947, *Les Cathodes Chaudes* (Paris: La Revue d'Optique théoret. instrum.), p. 127.

BLEWETT, J. P., 1939, *Phys. Rev.*, **55,** 713; 1946, *J. Appl. Phys.*, **17,** 643.

DE BOER, J. H., 1935, *Electron Emission and Adsorption Phenomena* (Cambridge: University Press).

COOMES, E. A., 1946, *J. Appl. Phys.*, **17,** 647.

EISENSTEIN, A., 1946, *J. Appl. Phys.*, **17,** 434; 1947, *Phys. Rev.*, **72,** 531.

EISENSTEIN, A., and FINEMAN, A., 1946, *J. Appl. Phys.*, **17,** 663.

FOWLER, R. H., 1936, *Statistical Mechanics* (Cambridge: University Press), p. 397.

LOWRY, E., 1930, *Phys. Rev.*, **35,** 1367.

MORGULIS, N. D., 1947, *J. Phys. Acad. Sci., U.S.S.R.*, **41,** 67.

NISHIBORI, E., and KAWAMURA, H., 1940, *Phys. Math. Soc., Japan*, **22,** 378.

REIMANN, A. L., 1934, *Thermionic Emission* (London: Chapman and Hall).

ROOKSBY, H. P., 1947, *Nature, Lond.*, **159,** 609. (Also 1940, *G.E.C. J.*, **11,** 83, and unpublished work.)

SEITZ, F., 1940, *Modern Theory of Solids* (New York and London: McGraw-Hill), p. 553.

SPROULL, R. L., 1945, *Phys. Rev.*, **67,** 166.

WEINREICH, M. O., 1945, *Gen. Elect. Rev.*, **54,** 243.

WRIGHT, D. A., 1947 a, *Proc. Roy. Soc.* A, **190,** 394; 1947 b, *Nature, Lond.*, **160,** 129; 1948, *Proc. Phys. Soc.*, **60,** 13, 22.

LETTERS TO THE EDITOR

Visual Adaptation and the Apparent Saturation of Colours

A binocular colour-matching apparatus has been used to investigate the way in which the appearance of coloured patches of light varies with the state of adaptation of the eye. Three adapting stimuli were used: the standard illuminant A (S_A), representing tungsten light, the standard illuminant B (S_B), representing daylight, and no stimulus, resulting in a state of dark-adaptation.

The right eye was kept dark-adapted throughout each experiment and viewed a $5° \times 2\frac{1}{2}°$ rectangular field of red, green and blue mixture of light, the colour and brightness of which could be controlled at will by the observer. The left eye viewed a $5° \times 2\frac{1}{2}°$ rectangular test patch of colour of luminance 2·3 candles/ft², which was kept of constant physical composition,* but which had a surround field of 60° which was either dark, or illuminated by S_A or by S_B at a luminance of 1·8 candles/ft².

The experimental procedure was first to match the test colour, say a pale yellow, as seen by the left eye, with a mixture of the red, green and blue light as seen by the right eye, with both eyes dark-adapted. The right eye was then kept dark-adapted while the left eye was light-adapted for five minutes to S_A. Two different arrangements were used in turn: first, the match was made in the presence of the adapting field surrounding the test patch; and secondly, the adapting field was blacked out during the moments when matching was actually being carried out. Then if under one of these conditions the test colour appeared to the left eye to be a deep orange colour, that fact would be recorded

* Except that, when the left eye surround field was dark, the test patch luminescence was lowered to 0·7 candles/ft² to facilitate matching.

quantitatively in terms of the amounts of red, green and blue light needed to produce an identical sensation in the dark-adapted right eye. Hence the dark-adapted right eye was used as a reference against which to measure changes in colour sensation due to adaptation in the left eye. The same procedure was then repeated using S_B as the adapting illuminant.

Some typical results are shown in Figure 1. The starting points of the full arrows represent the chromaticities of the various test colours used as measured when both eyes were dark-adapted. The tips of these arrows represent the chromaticities which it was necessary to present to the right eye for there to be a match when the left eye was light-adapted by S_B. Thus the changes of appearance of the test patch to the left eye, on changing its adaptation from dark adaptation to that of 1·8 candles/ft², of S_B, as recorded by a matching change in the dark-adapted right eye, are shown by the full arrows.

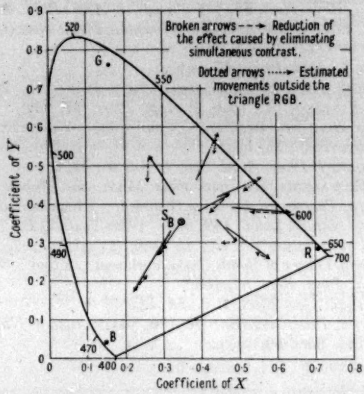


Figure 1. The movement of colours from their dark-adapted positions under the influence of S_B adaptation with surround.

It is seen that on light adaptation all the test colours become much more saturated in colour, and this increase in saturation was subjectively obvious to the observer and in many cases very striking.

The dotted arrows show the extent to which this effect is reduced by blacking out the adapting field during the actual time of matching. If the effect had been completely reduced by this procedure, we might have deduced that this increase of saturation was due to simultaneous contrast between the test patch and its white surround—conversely, since the reduction is only very partial, we can say that the saturation effect is not due mainly to simultaneous contrast.

Similar results were obtained when using S_A as the adapting illuminant in that a general increase in saturation was obtained, but there were important differences in the directions and the magnitudes of the arrows. These differences can be used to measure the difference between the sensations to which a test colour may give rise in an eye adapted to S_A or tungsten light and in an eye adapted to S_B or daylight. These differences are quite large and are of importance to industry in that full knowledge of them would make it possible to predict colour changes brought about by changing from daylight to tungsten light illumination, and vice versa.

It is of interest to see how results of the type shown in Figure 1 can be reconciled with our current view of a three-receptor retinal system. It is well known that colour adaptation can take place ; thus prolonged exposure to red light reduces the sensitivity of the red mechanism of the eye and results in vision in which the sensations are more blue-green than they would normally be. If the arrows of Figure 1 had all pointed away from the red corner of the diagram, such an explanation would of course be satisfactory ; but no amount of juggling with the overall sensitivities of the three receptor mechanisms could possibly result in arrows going simultaneously in all directions outwards from the centre.

Furthermore, trichromatic matches in which either one or both eyes view the test and comparison patches side by side remain matches over a very wide range of adaptation. This is usually regarded as implying that the spectral sensitivity curves of the three receptor mechanisms are of constant shape, although of course they change in overall height.

This would seem to indicate that the saturation effect of Figure 1 cannot be explained in terms of sensitivity changes of the receptor mechanisms. It is therefore suggested that the effect is due to interaction of the nerve impulses from the receptors somewhere between the retina and the brain. Such an explanation is shown schematically in Figure 2.

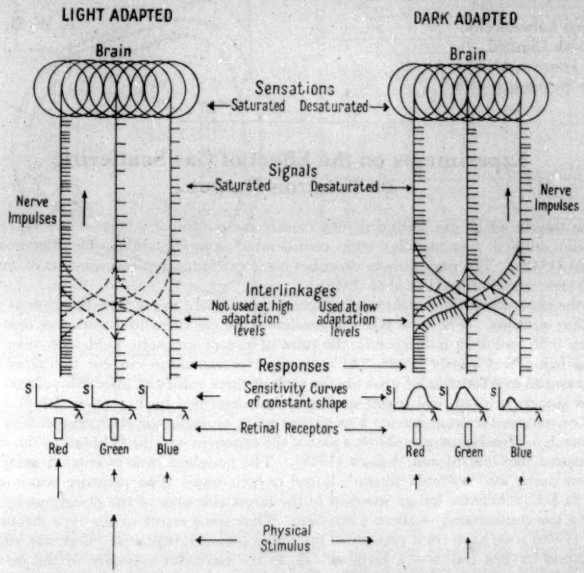


Figure 2. Schematic diagram of process proposed to explain the decrease in the saturation of sensations produced by the dark-adapted eye.

The left half of the diagram depicts the light-adapted state of affairs, while the right half depicts the dark-adapted. It is supposed that two physical stimuli, reddish in colour, of identical spectral distribution but different total energies, E and e, fall on the eye and that their magnitudes are roughly inversely proportional to the degree of adaptation, so that they appear of approximately equal brightness. The receptors of the light-adapted eye will have relatively low spectral sensitivity curves and those of the dark-adapted eye relatively high ones: the corresponding curves, however, have identical shapes, thus maintaining monocular and binocular trichromatic matches. In each case the three receptors produce their three responses, which are of the same order of magnitude, as shown. In the case of the light-adapted eye these three responses give rise directly to three signals to the brain which give a sensation which is not desaturated. In the case of the dark-adapted eye, however, the responses do not give three signals directly, but there is first a spilling over of the greatest response into the other two channels. These modified responses result in modified signals reaching the brain, and hence in this case a desaturated sensation is experienced.

It is felt that weight is added to these suggestions by the arborization of the retinal nerve paths between the receptors and the bipolar cells and between the bipolar cells and the ganglion cells of the retina, as reported by Polyak; and it seems natural to suppose

that the nerve impulses would intermingle more in the dark-adapted state than in the light-adapted, since greater sensitivity could be obtained in this way.

A full account of these investigations is being published in the *Journal of the Optical Society of America*.

The experimental part of this work was carried out at Imperial College, London, for Messrs. Kodak Limited. The author's grateful acknowledgments are due to Dr. W. D. Wright and Miss G. M. Frowde of Imperial College, to Dr. L. C. Thomson of the Vision Research Unit of the Medical Research Council, and to Mr. E. R. Davies and Mr. A. Marriage of Kodak Limited, who in various ways have helped him with this work.

Research Laboratories,
Kodak Limited,
Harrow, Middx.
9th November 1948.

R. W. G. HUNT.

Experiments on the Effect of Gas Scattering on Betatron Output

The way in which gas scattering may modify the output of a betatron or synchrotron has been studied theoretically, with considerable approximations, by Blachman and Courant (1948). The present note describes some preliminary measurements of the effect in a 30 MEV. synchrotron (Fry *et al.* 1948).

In the experiment the synchrotron magnet was operated as a 4 MEV. betatron at various excitation currents. The time taken to accelerate the electrons to 4 MEV. was thus varied between 0·56 and 2·56 milliseconds, the rates of rise of magnetic field with time in this interval being very nearly linear. At each value of excitation current the X-ray output was measured as a function of pressure for each of three values of injection voltage.

The doughnut used, which was similar to that described by Fry *et al.* (1948), had the resonator removed from it, leaving a minimum space available for electron orbits as shown in Figure 1. This Figure also shows a plot of the exponent n in the field law of the magnet, as measured by Goward and Wilkins (1948). The doughnut was evacuated using an oil diffusion pump and a " cold finger " liquid oxygen trap. The pressure was measured with a G.E.C. ionization gauge attached to the target side-arm of the doughnut by a tube having a gas conductance of about $\frac{1}{2}$ litre/sec. Thus some errors of the type discussed by Blears (1946) may have been present in the lower pressure readings. Gas was admitted, via a liquid oxygen trap and a variable leak, to the resonator side-arm of the doughnut.

It was found that graphs of log (output) plotted against pressure closely approximated to straight lines, the slope of which depended on the acceleration period of the electrons and on the injection voltage. Specimen curves plotted for air are shown in Figure 2. These graphs may be fitted to the law

$$I = I_0 \exp\left(-Ptf(V)\right), \qquad \ldots\ldots (1)$$

where I is the X-ray intensity at a pressure P millimicrons of mercury, I_0 is the intensity at zero pressure, and t is the time required to accelerate the electrons to 4 MEV. measured in milliseconds. $f(V)$ is some function of the injection voltage V, measured in kilovolts, and has the values shown in the Table as determined from graphs similar to Figure 2:

V (in kv.)	·30	30	30	30	16	16	16	16	10	10	10
t (in msec.)	2·46	1·14	0·75	0·56	2·46	1·14	0·75	0·56	2·46	1·14	0·75
$f(V)$	0·17	0·23	0·23	0·19	0·24	0·30	0·30	0·25	0·28	0·34	0·39

These values are in fair agreement with an empirical relation $f(V) = 1{\cdot}1\, V^{-\frac{1}{2}}$, and equation (1) thus becomes

$$I = I_0 \exp\left(-1{\cdot}1\, PtV^{-\frac{1}{2}}\right). \qquad \ldots\ldots (2)$$

On introducing hydrogen into the doughnut (but using the same calibration of the ionization gauge as for air), the variation of output with gas pressure was considerably reduced, equation (2) becoming

$$I = I_0 \exp\left(-0{\cdot}09\, PtV^{-\frac{1}{2}}\right). \qquad \ldots\ldots (3)$$

Eddy currents in the magnet cause some inaccuracy in the above results by altering the effective cross-section of the doughnut and thus affecting the slope of the output–pressure curves. Their presence was demonstrated by the rapid decrease, with increasing magnet excitation, of the zero-pressure output of the machine when operated at the lower injection voltages.

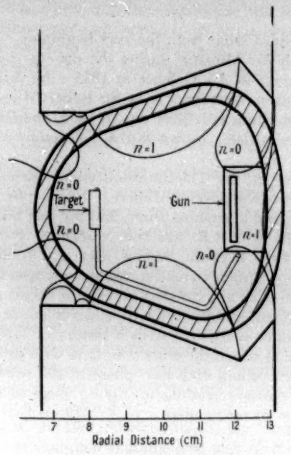


Figure 1. Doughnut and magnetic field of 30 MEV. synchrotron.

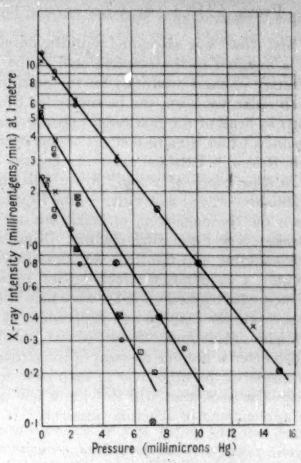


Figure 2. Graph of log output *V*. Pressure for air.
Magnet excitation 40 amp.
Upper, middle and lower curves: injection voltages 30, 16, 10 kv. respectively.

In the practical operation of a betatron or synchrotron there is often doubt as to what gases are present in the doughnut. Thus, using an earlier type of doughnut internally coated with graphite, Dain and Goward (1948) obtained outputs which depended on the ultimate pressure attained, in a manner approximately described by equation (2). The pressure rose rapidly during operation, however, due to the liberation of gases from the graphite by heat and electron bombardment, and these gases affected the output in a manner much more nearly described by equation (3). The doughnut used in the present experiments had a fired-on platinum coating, which did not evolve any gas under operating conditions. Thus the variations of output with pressure were solely due to the deliberately introduced gas.

The results of this note give suitable criteria for the maximum gas pressure permissible in a betatron or synchrotron. Further, the exponential form of the output–pressure curves demonstrates the absence of any appreciable gas focusing of the electron beam.

This note is published by permission of the Director of the A.E.R.E., Harwell.

Atomic Energy Research Establishment,
Harwell.
1st December 1948.

H. H. H. WATSON.

BLACHMAN, N. M., and COURANT, E. D., 1948, *Phys. Rev.*, **74,** 140.
BLEARS, J., 1946, *Proc. Roy. Soc.* A, **188,** 62.
DAIN, J., and GOWARD, F. K., 1948, private communication.
FRY, D. W., GALLOP, J. W., GOWARD, F. K., and DAIN, J., 1948, *Nature, Lond.*, **161,** 504.
GOWARD, F. K., and WILKINS, J. J., 1948, *Proc. Phys. Soc.*, **61,** 580.

REVIEWS OF BOOKS

One Story of Radar, by A. P. Rowe. Pp. xii + 208. (Cambridge: University Press, 1948.) 8*s*. 6*d*.

Mr. Rowe was associated with the development of radar from the very beginning. He was a member of the research staff of the British Air Ministry during the pre-war years, Secretary of the Committee for the Scientific Survey of Air Defence in 1935—the year in which radar was born, and he succeeded R. A. Watson Watt (later to be Sir Robert Watson-Watt) as head of the first radar experimental establishment. This was in 1938, and Rowe remained Chief Superintendent of the establishment (later known as the Telecommunications Research Establishment) during the war years.

In *One Story of Radar*, Rowe tells the story of the Telecommunications Research Establishment, and how radar for the Royal Air Force was developed there. It is not by any means the complete story of radar, but one story, and a personal story, and for that reason it makes very interesting reading. Rowe's theme at T.R.E. was the creation of a close understanding and informal yet constant cooperation between Service user and scientist inventor, a cooperation which proved to be one of Britain's greatest assets. He tells how the men who worked on radar were men with a mission, and his story reveals the excitement of the work, the elations and depressions as it faltered, and the pride in achievement of the men who did the job. Rowe, in fact, has done much more than write a history of T.R.E., he has given a graphic account of the atmosphere in which the work was done with proper discussion of the objectives of each new development and each new phase of the work.

It is a book that will provide a few hours' pleasant reading to anyone interested in the applications of scientific research. It can be recommended. DENIS TAYLOR.

Fundamentals of Vibration Study, by R. G. Manley. Pp. xvi + 158. Second Edition. (London: Chapman & Hall, Ltd., 1948.) 15*s*.

The study of engineering vibrations has had its circle of specialists and devotees for many years, but only in recent years has the subject become of such importance that no engineer can ignore it. The advent of high speed prime movers and machines has demanded much greater attention to unnecessary mass in machine parts. The large factors of safety which have allowed the designing engineer to ignore the effects of cyclical stresses must be reduced to a minimum and that failure by fatigue mainly controls the designer's considerations.

Accordingly the engineer and the physicist need an authoritative presentation of the elements of vibrations, and the fact that a second edition of a book first published in 1942 has already been called for would indicate that Mr. Manley has gone a long way to supplying that need.

The first chapters deal with simple vibrating systems having one degree of freedom, both in undamped and damped motion, followed by the effects of superimposed forcing vibrations upon the characteristics of such systems.

These considerations lead to systems having two degrees of freedom, and to complicated systems having more than two degrees of freedom. In this section it is shown that the use of the concepts of effective inertia and dynamic stiffness facilitates the determination of natural frequencies.

Continuous systems—such as vibrating beams and shafts—are dealt with in the next chapter, while the last chapter introduces the subject of vibrations of non-sinusoidal or irregular shape.

The book ends with five lengthy appendices. In some cases it would appear that these could have been profitably included in the text, but perhaps publication difficulties prevented the appropriate sections being rewritten.

The treatment of some of the subject matter is uneven. For example, in Appendix II Newton's second law of motion is dealt with, together with the elementary formulae for the bending of beams. These matters are widely known and available in standard

textbooks. On the other hand the properties of the differential operator D are not stated: this notation, and the distributive, commutative etc., laws of the operator, are not so well known. Half a page in explanation would have been sufficient to refresh the memory. Again, to anyone studying the subject for the first time, the presence of the negative sign in the equation $m\ddot{x} = -kx$ could very well have been explained. The general approach in the first chapter might have been rendered easier by a slight rearrangement.

It is a pity that the engineer is still supposed to need his mass unit in slugs: a clear statement of fundamental units obviates the use of a unit which is more likely to confuse him than help him. The slug, as defined in this book, in terms of inches per sec. per sec. is an archaism at which even the engineer shudders. HUGH FORD.

Velocity-Modulated Thermionic Tubes, by A. H. W. BECK. Pp. x+180. (Cambridge: University Press, 1948.) 15*s*.

This book provides a very useful introduction to velocity-modulation thermionic tubes and their mode of operation and should be readily intelligible to anyone with a reasonably adequate knowledge of pre-1939 radio technique. The book contains a short discussion on resonant cavities and on heavy current electron beams, but the major part of the book is concerned with the general theory of the interchange of energy between field and beam and the application of this theory to the various types of velocity-modulation tubes. Chapters are included on velocity-modulation amplifiers and frequency multipliers, klystron oscillators and reflex klystron oscillators and reference is made to a number of special tubes.

Little is said about the uses to which velocity-modulated tubes are put or about measurements on them, but the author has attempted successfully to build up a theory from simple postulates by successive refinements so that physical facts dominate mathematical conditions, and a realistic picture is obtained of the way in which a physical theory of the V.M. tube is arrived at.

The last two chapters discuss the limitations and defects in V.M. tubes and the design requirements and manufacture of V.M. tube components, and there are two appendices—one outlining the design calculations of a high power C.W. klystron and the other discussing the travelling wave tube of Kampfer.

This is the fourth volume of the Modern Radio Technique Series edited by Mr. J. A. Ratcliffe, and it will certainly prove to be a valuable companion volume to these books. It can be strongly recommended. DENIS TAYLOR.

Tables of Physical and Chemical Constants, by G. W. C. KAYE and T. H. LABY. Pp. vii+194. (London, New York and Toronto: Longmans, Green & Co., 1948.) 21*s*. net.

This tenth edition of "Kaye and Laby" was prepared in Melbourne, presumably on account of the death of Dr. Kaye during the preparation of the ninth edition. It has thus lost some of its associations with the National Physical Laboratory, but it has been fortunate in gaining new assistance in its compilation.

Considerable revision has been undertaken, and in particular the fundamental constants have been corrected in the light of recent determinations and derived constants have been accordingly re-calculated. Similarly, the astronomical constants have been revised, and the Astronomer Royal's determination of the solar parallax is adopted. A useful section on the properties and constants of optical glasses has been added.

A glance at selected tables failed to disclose any misprints, and the general standard of production of the book is very high. H. H. HOPKINS.

Bulletin Analytique. Vol. VIII, No. 11–12, Parts I and II. (Paris: Hermann et Cie, 1947.) Part I, pp. 2577–2964; Part II, pp. 2025–2345.

What *Science Abstracts* attempts to do for Physics, and *Chemical Abstracts* for Chemistry, the *Bulletin Analytique* undertakes for the whole range of Science—Physical and Biological, Pure and Applied. It takes nothing less than the whole of scientific literature for its domain. This colossal venture is in the hands of a sub-section of the *Centre National*

de la Recherche scientifique known as the *Centre de Documentation*, a special branch of which carries out the actual editorial duties. The cover also bears the imprint of the Ministère de l'Éducation Nationale.

Of the two volumes of the double number which have come to hand, the first (1re Partie) deals with the physical sciences and their applications, the second with the biological sciences and their ancillary techniques. One imagines, from the pagination, that each "Partie" is intended to form its own separate volume at the end of the year. The two parts together comprise rather more than 700 pages, and include more than 8,000 abstracts in all. It is a sobering thought that, on the evidence of this journal, scientific papers are being turned out at the rate of one every ten minutes, night and day, day in and day out!

In style and length the individual abstracts are similar to those with which we are familiar in *Science Abstracts*: adequate to indicate the scope of the publication, but not to absolve the reader from the necessity of consulting it. In this connection, however, the Centre de Documentation offer to supply readers with a microfilm reproduction of any of the papers abstracted: a scheme which may be commended to the attention of the learned Societies responsible for the publication of our own *Science Abstracts*.

One gathers that the *Bulletin Analytique* culls its extracts from a rather wider field than our own abstracting journals, and that, in addition to the standard scientific journals, technical, semi-popular and even trade journals are scanned for matter worthy of note. One learns, for example, of the part played by amateur radio fans during the Texas disasters; that geologists have made an excursion to Reading; that Plastics has been looking at Brewing, and that, in view of the recent drought, Valparaiso is to accelerate its plans for a new water supply.

A very valuable addition to the normal abstracting service is the list of new books of scientific, and semi-scientific interest, which appears at the end of each Part of the *Bulletin*. An author index is also included with each part. Our French colleagues have no cause to complain of the efficiency of their abstracting service. J. A. CROWTHER.

Physico-Chemical Methods, Vol. III, by J. REILLY and W. N. RAE. Pp. ix + 697. (London: Methuen & Co. Ltd., 1948.) 55*s*.

The research worker and advanced student interested in the border line problems of Chemistry and Physics has particular need of a handbook which will provide an authoritative guide to the experimental methods available in this very extensive field. The provision of such a book is no easy task in view of the great variety of methods used, many of which involve highly specialized and elaborate techniques, and it is very difficult to secure adequate treatment of each topic and yet to keep the entire work within reasonable compass. Reilly and Rae's *Physico-Chemical Methods*, first published in 1926, has long been valued in this connection, and the fourth revised edition was published in two volumes in 1943. The present volume is intended as a supplement to this last edition. It contains accounts of a number of aspects of the subject not treated in the previous volumes, such as electron optics, mass spectrography, and the determination of atomic weights, while in other sections, such as the determination of molecular weights, reaction kinetics, chromatography, high vacuum technique, colloids, electrolytic oxidation and gas analysis, the earlier accounts have been supplemented and in some cases completely re-written. The general level of treatment of the various topics is high and the information given is full and up to date. While the expert will probably find points to criticize in the account of his own special field, the book is a mine of information about methods and apparatus, and the extensive bibliography provided indicates where further details can be found. The present work seems likely to join the two earlier volumes as an indispensable tool for the practical worker in Physical Chemistry or Chemical Physics. As is perhaps inevitable in a supplement which is intended partly to add to, and partly to replace, earlier accounts, the present volume is not very satisfactorily integrated with the existing volumes I and II, but an early revision of these is promised. The book is well printed, lavishly provided with diagrams, and of convenient format, and the price, although high, is not unreasonable for a work of this type. A. H.

CONTENTS FOR SECTION A

PAGE

Mr. E. Bauer. A Theory of Ultrasonic Absorption in Unassociated Liquids . 141

Mr. R. B. Dingle. Second Sound and the Behaviour of Helium II . . . 154

Dr. R. Shuttleworth. The Surface Energies of Inert-gas and Ionic Crystals 167

Dr. K. Huang and Mr. G. Wyllie. On the Surface Free Energy of Certain Metals 180

Dr. R. C. Pankhurst. The Emission Spectrum of Sodium Hydride. . . 191

Dr. L. Kellner. Bending Vibrations of a Linear Chain 200

Letters to the Editor :

Dr. F. C. Frank. Sessile Dislocations 202

Contents for Section B 203

Abstracts for Section B 203

ABSTRACTS FOR SECTION A

A Theory of Ultrasonic Absorption in Unassociated Liquids, by E. Bauer.

ABSTRACT. In many liquids the absorption of sound waves is very much larger than can be explained classically, in terms of viscosity and heat conduction. It is now fairly well established that different types of liquids behave very differently in this respect, and that the cause of the excess absorption in unassociated, polyatomic liquids like benzene and carbon disulphide lies in the incomplete excitation, by the sound wave, of the vibrational degrees of freedom. It is shown that this effect leads to mean dispersion frequencies in the range 10^9 to 10^{10} c/s. at 300° K. This frequency range shows that, as in the case of gases, the efficiency of collisions in producing excitation or de-excitation of vibrational degrees of freedom is fairly low; a theoretical treatment of this effect is outlined. Detailed considerations of temperature coefficients of absorption lead to the conclusion that all the vibrational levels of a liquid do not relax at the same frequency but that high excited levels relax at relatively low frequencies. In the case of carbon disulphide a numerical estimate of these individual dispersion frequencies is made. It is also shown that the variation of absorption with concentration in binary mixtures can be described quite simply in terms of the analysis developed in this paper.

Second Sound and the Behaviour of Helium II, by R. B. Dingle.

ABSTRACT. In this paper some theoretical consequences of the existence of second sound in helium II are discussed. A method is suggested for setting up a quantitative phenomenological description of the thermodynamical behaviour of helium II. It is known that expressions for the velocities of first and second sound may be obtained by using quite general models. The thermodynamical behaviour may be described in terms of the velocities of these two types of sound, so that it becomes possible, in principle, to obtain explicit values of the unknowns by combining the two sets of equations. On the basis of these ideas a discussion is given of the coefficient of thermal expansion of helium, the ordinary thermal conductivity, and the attenuation of second sound.

The paper contains a discussion of the modifications necessary to the Debye theory of specific heats when the sound velocities are temperature-dependent.

The Surface Energies of Inert-Gas and Ionic Crystals, by R. SHUTTLEWORTH.

ABSTRACT. It is shown that the surface energies of inert-gas crystals are proportioned to their sublimation energies. The constant of proportionality is evaluated for the {111}, {100} and {110} faces for different assumed forms of the interatomic repulsive forces; to within 15% it is independent of the face, and of the form of the repulsive force. The distance of the outermost plane of atoms from the next at a {100} face is 2·54% greater than normal; this distortion makes a negative contribution of 1% to the surface energy.

The surface energies of sodium and potassium halide crystals are calculated; the ratio of the energies of {110} and {100} faces is about 2·5. The contribution of van der Waals type forces to the surface energy of a {100} face is almost half the total.

Elementary methods are given for the evaluation of lattice sums.

On the Surface Free Energy of Certain Metals, by K. HUANG and G. WYLLIE.

ABSTRACT. For the purpose of calculating the surface energy and surface double layer for a number of monovalent metals, a simple and yet adequate model is proposed, which does not have the disadvantages of models employed by previous authors. In this the valence electrons are treated according to a simple Sommerfeld model, and the positive charge of the ions is distributed uniformly through the metal. It is shown that by a suitable combination of the methods of Thomas-Fermi and of direct wave mechanics, both the surface energy and the strength of the double layer can be calculated. The surface double layers for the metals considered are found to lie in the range 0·1–0·5 ev. The contribution to the surface energy from the electrostatic energy in the double layer is found to be very small. In estimating the surface energy, the electrons in a metal can be considered as moving in a box with finite potential walls of suitably determined height ϕ. A temperature-dependent term in the surface free energy due to the change in lattice vibrations caused by the free surface is also considered. The surface energies obtained from ϕ determined in two independent ways agree closely and are found to agree with the experimental values of the surface tension of the liquid metals to within a factor 1·5.

The Emission Spectrum of Sodium Hydride, by R. C. PANKHURST.

ABSTRACT. Observations on the spectrum of sodium hydride have been extended to longer wavelengths (from 4600 A. to 6450 A.) and photographed under high dispersion. Analysis of the data thus obtained gives values of the vibrational and rotational constants for several new vibrational levels, among which the lowest levels of the excited state are of especial interest.

The intensity distribution has been calculated from the wave functions appropriate to potential energy curves based on the constants derived from the analysis of the band system. The theoretical estimates are found to be in fair agreement with the experimental results.

Bending Vibrations of a Linear Chain, by L. KELLNER.

ABSTRACT. The bending frequencies of a linear chain of n mass points have been calculated using a harmonic potential function involving forces between second neighbours. The proper boundary conditions are introduced and a general formula for the bending frequencies is derived.